A SPECIAL FAMILY OF BINARY FORMS, THEIR INVARIANT THEORY, AND RELATED COMPUTATIONS.

by

CHRISTOPHE DETHIER

A DISSERTATION

Presented to the Department of Mathematics
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2020

DISSERTATION APPROVAL PAGE

Student: Christophe Dethier

Title: A Special Family of Binary Forms, Their Invariant Theory, and Related Computations.

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Mathematics by:

Shabnam Akhtari	Chair
Nicholas Addington	Core Member
Ellen Eischen	Core Member
Benjamin Young	Core Member
Hank Childs	Institutional Representative

and

Kate Mondloch	Interim Vice Provost and Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded June 2020

DISSERTATION ABSTRACT

Christophe Dethier

Doctor of Philosophy

Department of Mathematics

June 2020

Title: A Special Family of Binary Forms, Their Invariant Theory, and Related Computations.

In this manuscript we study the family of diagonalizable forms, a special family of integral binary forms. We begin with a summary of definitions and known results relevant to binary forms, diagonalizable forms, Thue equations, and reduction theory.

The Thue–Siegel method is applied to derive an upper bound on the number of solutions to Thue's equation $F(x, y) = 1$, where F is a quartic diagonalizable form with negative discriminant. Computation is used in the argument to handle forms whose discriminant is small in absolute value. These results are applied to bound the number of integral points on a certain family of elliptic curves.

A proof is given for an alternative classification of diagonalizable forms using the Hessian determinant. Algebraic restrictions are given on the coefficients of a diagonalizable form and divisibility conditions are given on its discriminant. A reduction theory for the family of diagonalizable forms is given. This theory is used to computationally verify that $F(x, y) = 1$, where F is a quintic diagonalizable form with small discriminant, has few solutions.

CURRICULUM VITAE

NAME OF AUTHOR: Christophe Dethier

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, OR
Carleton College, Northfield, MN

DEGREES AWARDED:

Doctor of Philosophy in Mathematics, 2020, University of Oregon
Bachelor of Arts in Mathematics, 2013, Carleton College

AREAS OF SPECIAL INTEREST:

Number Theory
Diophantine Approximation
Computational Mathematics

PROFESSIONAL EXPERIENCE:

Graduate Teaching Fellow, University of Oregon, September 2013 - June 2020

PUBLICATIONS:

C. Dethier. Diagonalizable quartic Thue equations with negative discriminant *Acta Arithmetica* 193: 235–252, 2020

ACKNOWLEDGEMENTS

I thank Professor Akhtari with her advice and guidance throughout my doctoral program. I thank the members of my committee for their advice in preparing this manuscript. I also thank Professor Sinclair for his advice and mentorship.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

1.1. Binary Forms

A binary form of degree r is a homogenous polynomial of degree r in two variables, perhaps

$$F(x,y) = a_0x^r + a_1x^{r-1}y + \ldots + a_{r-1}xy^{r-1} + a_ry^r,$$

with $a_0, a_1, \ldots, a_{r-1}, a_r \in \mathbb{C}$. In this manuscript we are particularly interested in integral binary forms, those binary forms for which $a_0, a_1, \ldots, a_{r-1}, a_r \in \mathbb{Z}$. For each binary form $F(x,y)$ we associate a univariate polynomial $f(x) = F(x,1)$, and for each univariate polynomial $f(x)$ we associate a binary form $y^r f(x/y)$. These associations are inverse and can be used to translate many definitions for univariate polynomials to binary forms. The following two definitions are examples of this.

A binary form F splits into linear forms over the complex number as the corresponding univariate polynomial $f(x) = F(x,1)$ splits over the complex numbers. Thus F can be written as

$$F(x,y) = (x - \alpha_1 y)(x - \alpha_2 y)\ldots(x - \alpha_r y),$$

and these $\alpha_1, \ldots, \alpha_n$ are called the roots of F. The roots of F are the exactly the roots of $f(x) = F(x,1)$. The discriminant of F is then given by

$$\Delta_F = a_0^{2r-2}\prod_{i<j}(\alpha_i - \alpha_j)^2,$$

where a_0 is the leading coefficient of F, as above. As the roots of $F(x, y)$ are exactly the roots of $f(x) = F(x, 1)$, the discriminant of F is exactly the discriminant of $f(x) = F(x, 1)$.

An integral binary form F is said to be irreducible if there is no decomposition of F as $F(x, y) = G(x, y)H(x, y)$, where G and H are integral binary forms of degree at least one. A binary integral form is irreducible over $\mathbb{Z}$ exactly when the associated univariate polynomial $F(x, 1)$ is irreducible over $\mathbb{Z}$.

Suppose that F is an integral binary form and $\mathcal{S}$ is a subgroup $\mathrm{GL}_2(\mathbb{C})$. If the matrix

$$\begin{pmatrix} a_{00} & a_{10} \\ a_{01} & a_{11} \end{pmatrix} \text{ is in } \mathcal{S},$$

then we say that the form

$$G(x, y) = F(a_{00}x + a_{10}y, a_{01}x + a_{11}y)$$

is an $\mathcal{S}$-substitution of F, or that the forms F and G differ by an $\mathcal{S}$-substitution. In the special cases $\mathcal{S} = \mathrm{GL}_2(\mathbb{Z})$ and $\mathcal{S} = \mathrm{SL}_2(\mathbb{Z})$, we say that F and G are equivalent or properly equivalent respectively.

An invariant $\mathcal{I}$ is a homogeneous integral polynomial in the coefficients of a binary form which changes by a determinental factor under $\mathrm{GL}_2(\mathbb{C})$ substitution. That is, if $a_{00}, a_{10}, a_{01}, a_{11} \in \mathbb{C}$ with $a_{00}a_{11} - a_{10}a_{01} \neq 0$, and $G(x, y) = F(a_{00}x + a_{10}y, a_{01}x + a_{11}y)$, then $\mathcal{I}$ satisfies

$$\mathcal{I}(G) = (a_{00}a_{11} - a_{10}a_{01})^k \mathcal{I}(F) \tag{1.1}$$

for some positive integer k, called the weight of $\mathcal{I}$. The ring of such invariants was famously shown to be finitely generated by Hilbert in his celebrated basis Theorem.

Generators for the invariant ring are known in small degree. The quadratic and cubic invariant rings are generated by the quadratic and cubic discriminants respectively. The quartic invariant ring is generated by two invariants, I and J of weight 2 and 3 respectively. If $F(x, y)$ has coefficients

$$F(x, y) = a_0x^4 + a_1x^3y + a_2x^2y^2 + a_3xy^3 + a_4y^4,$$

then

$$I_F = a_2^2 - 3a_1a_3 + 12a_0a_4$$

and

$$J_F = 2a_2^3 - 9a_1a_2a_3 + 27a_1^2a_4 - 72a_0a_2a_4 + 27a_0a_3^3.$$

The invariant ring for binary quintics is generated by the four invariants with one relation. Not even the number of generators is known in high degree.

We conclude with an important definition. The Hessian determinant of F is given by

$$H(x, y) = H_F(x, y) = \left(\frac{\partial^2 F}{\partial x^2}\right)\left(\frac{\partial^2 F}{\partial y^2}\right) - \left(\frac{\partial^2 F}{\partial xy}\right)^2.$$

1.2. Diagonalizable Forms

A diagonalizable form of degree r is an integral binary form which has the shape

$$F(x, y) = (\alpha x + \beta y)^r - (\gamma x + \delta y)^r \tag{1.2}$$

for some $\alpha, \beta, \gamma, \delta \in \mathbb{C}$ with

$$j = \alpha\delta - \beta\gamma \neq 0.$$

Furthermore, there is a constant χ such that

$$(\alpha x + \beta y)(\gamma x + \delta y) = \chi(Ax^2 + Bxy + Cy^2).$$

The linear forms $u(x, y) = \alpha x + \beta y$ and $v(x, y) = \gamma x + \delta y$ are sometimes referred to as the resolvent forms of F.

Sylvester's canonical forms given in [1] and [2] show that every quadratic and cubic form is diagonalizable. For this reason, we restrict ourselves to $r \geq 4$.

We turn to the question of how one might determine whether an arbitrary integral binary form is diagonalizable, and if so, what resolvent forms it may be constructed from. A general method for answering the first question is through Gundelfinger's result, proved in [3], which has, as a special case, that F is diagonalizable if and only if $G_2[F] \equiv 0$. Here $G_2[F]$ is the second Gundelfinger covariant, which is the 3×3 determinant

$$G_2[F] = \det\left[\left(\frac{\partial^4 F}{\partial x^{4-i-j} y^{i+j}}\right)_{0 \leq i,j \leq 2}\right].$$

Gundelfinger's result in full generality allows one to determine when a form can be expressed as a sum of a fixed number of rth powers of linear forms.

Although Gundelfinger's result is sufficient for practical purposes, we pursue this issue further. One should be able to describe restrictions like diagonalizability on the shape of the form in terms of the vanishing of certain generators of the invariant ring. For example, a quartic binary integral form F is diagonalizable if

and only if $J_F = 0$. This result is stated in [4, p. 29], and shown explicitly in [5]. If the Hessian H_F of F is written

$$H_F = \frac{\partial^2 F}{\partial x^2}\frac{\partial^2 F}{\partial y^2} - \left(\frac{\partial^2 F}{\partial x \partial y}\right)^2 = A_0x^4 + A_1x^3y + A_2x^2y^2 + A_3xy^3 + A_4y^4,$$

then in [5] it was shown that

$$F(x,y) = \frac{1}{8\sqrt{3I_FA_4}}\left(\xi^4(x,y) - \eta^4(x,y)\right), \tag{1.3}$$

where ξ^4 and η^4 have coefficients in $\mathbb{Q}\left(\sqrt{A_0I_F/3}\right)$. Furthermore if $I_F > 0$ then ξ and η are complex conjugates.

Every diagonalizable form is determined by a $\mathrm{GL}_2(\mathbb{C})$-substitution of the form $x^r - y^r$. We see from (1.1) that the invariants which vanish are not altered by such substitutions. Thus we may evaluate the generators of the invariant ring for the form $x^r - y^r$ to see that, for example, every quintic diagonalizable form has the same three generating invariants vanish. Although we suspect that this vanishing is a sufficient condition for diagonalizability, it has not been shown. One can perform similar computations in every degree for which generators of the invariant ring are known.

For the second question, determining the possible values of α, β, γ, and δ in the diagonalization (1.2) of F, we turn to the Hessian. Computing the Hessian of (1.2), we see that

$$H_F(x,y) = r(r-1)j^2((\alpha x + \beta y)(\gamma x + \delta y))^{r-2}. \tag{1.4}$$

Thus one can determine the resolvent forms of a diagonalizable form by factoring the Hessian. Interestingly, the converse of this computation also holds.

Theorem 1.2.1. *Suppose that $F(x, y) \in \mathbb{Z}[x, y]$ is an integral binary form of degree r with nonzero discriminant. Then F is diagonalizable if and only if H_F is the $r-2$ power of a quadratic form with non-proportional linear factors.*

1.3. Thue Equations

Suppose that $F(x, y)$ is a binary integral form whose irreducible factors are of degree at least three, and $h \in \mathbb{Z}$ is nonzero. Thue proved in [6] that the equation

$$F(x, y) = h \tag{1.5}$$

has finitely many solutions $(x, y) \in \mathbb{Z}^2$. Such equations are called Thue equations. It follows that inequalities of the form

$$0 < |F(x, y)| \leq h \tag{1.6}$$

also have finitely many solutions $(x, y) \in \mathbb{Z}^2$. Such inequalities are referred to as Thue inequalities. We note that if F in any of these equations is replaced by an equivalent form, then the number of solutions does not change.

Although Thue proved finiteness, giving bounds on the sizes of the solutions or on the number of solutions is of particular interest. In Chapter II we pursue the latter for a specific family of forms. To do this, we use the Thue–Siegel method of approximating binomial functions using Padé approximation. This method was developed by Thue, see for example [7]. The approximating functions were identified by Siegel as hypergeometric functions in [8]. The specifics of our

application of this method are derived from the work of Akhtari, Saradha, and Sharma found in [9].

A primitive solution to Thue's equation or inequality is a solution (x, y) for which $x \geq 0$ and $\gcd(x, y) = 1$. Throughout this manuscript we only count primitive solutions.

Akhtari applied the Thue–Siegel method in [5] to show that $|F(x, y)| = 1$ has at most 12 solutions when F is a diagonalizable quartic form with positive discriminant. In [10], Akhtari gives further results concerning the diagonalizable case with positive discriminant. In [11], Siegel shows that $0 < |F(x, y)| \leq h$, where F is diagonalizable quartic with negative discriminant, has at most 16 solutions when $D < 0$, at most 8 solutions when $D > 0$ and F is indefinite, and at most 1 solution when $D > 0$ and F is definite, all provided that $|\Delta_F| > 2^{59}h^{13}$.

Akhtari, Saradha, and Sharma applied similar methods in [9] to give similar bounds on the number of solutions to $|F(x, y)| = 1$ when F is diagonalizable of degree at least five. Quartic Thue inequalities have been studied by others, notably Wakabayashi in [12] and [13].

Chapter II concerns the case when F is diagionalizable quartic with negative discriminant using the methods of [5] and [9]. Using gap principles from [9] we prove that $0 < |F(x, y)| \leq h$ has at most $2k$ solutions under roughly the condition $h \ll_k 2^{-10/7}|j|^{10/7}$.

Theorem 1.3.1. *Let F be a diagonalizable quartic form with negative discriminant, and k an integer satisfying $k \geq 3$. Suppose that $h < \frac{1}{4}|j|^2$ and $h < C_2(2, k, 0)|j|^{E_2(2,k,0)}$, where*

$$E_2(2, k, 0) = \frac{110 \cdot 3^k - 1278}{77 \cdot 3^k + 378}$$

and $C_2(2,k,0) = 2^{\Theta}$*, where*

$$\Theta = \frac{108\log_2(3) - 6066 - 110\cdot 3^k}{378 + 77\cdot 3^k}.$$

Given these assumptions on h*, the Thue inequality* $0 < |F(x,y)| \leq h$ *has at most* $2k$ *primitive solutions.*

We refer to the exposition preceeding Lemma 2.3.1 for the complete definition of $C_2(n,k,g)$ and $E_2(n,k,g)$ where n, k, and g are integers satisfying $n \geq 2$, $k \geq 3$, and $g = 0, 1$.

Applying Theorem 1.3.1 in the case when $h = 1$ and $k = 4$ yields the following:

Theorem 1.3.2. *Let* F *be a diagonalizable binary quartic form with negative discriminant. The equation* $|F(x,y)| = 1$ *has at most eight primitive solutions.*

Our method of proof for Theorem 1.3.2 is to use Theorem 1.3.1 when $h = 1$. However this does not apply to forms with small $|\Delta|$, so we compute the solutions to $|F(x,y)| = 1$ for the remaining forms. Using $k = 4$ instead of $k = 3$ results in a more feasible computational problem. We refer the reader to Section 2.5 for the details of the computational methods used and some remarks on the results of these computations.

Diagonalizable forms are useful because if one can give an upper bound on the number of solutions to the Thue equation (1.5) with $h = 1$ and F diagonalizable, then one can give an upper bound on the number of solutions to the equation (1.5) with $h \in \mathbb{Z}$ nonzero and F is diagonalizable using a reduction of Bombieri and Schmidt found in [14]. See Proposition 2.7.1 for our specific version

of this. If given a diagonal form, that is one of type

$$F(x, y) = ax^n - by^n, \tag{1.7}$$

the Bombieri–Schmidt reduction will not necessarily return diagonal forms, but will return diagonalizable forms.

Applying the Bombieri–Schmidt reduction to Theorem 1.3.2 gives the following result:

Theorem 1.3.3. *Let G be a diagonalizable quartic form with negative discriminant. Then $|G(x, y)| = h$ has at most $8 \cdot 4^{\omega(h)}$ primitive solutions.*

We finish this chapter by applying this result to give an upper bound on the number of integral points on the elliptic curve

$$Y^2 = X^3 + NX \tag{1.8}$$

where N is a positive integer. We use the reduction found in [15]. In that paper, Tzanakis uses norm-form equations to give a method of finding the integral points on (1.8) but does not give an explicit upper bound on the number of such points. Tzanakis also gives a reduction for the same family of elliptic curves with N a negative integer (corresponding to a positive discriminant of the resulting forms), which Akhtari applied in [10] using the results from [5]. We have shown the following result using these methods:

Theorem 1.3.4. *Let N be a positive square-free integer. The equation* (1.8) *has at most*

$$2^{15/2}\sqrt{N}\sum_{d|N}\frac{2^{\omega(N/d)}\epsilon_d^{3/2}}{d}$$

integral points, where ϵ_d is a minimal unit in the ring $\mathbb{Z}[\sqrt{d}]$.

Reducing questions about integral points on an elliptic curve to solving a number of quartic Thue equations is a classical idea. See [16] for a recent computational example which uses the correspondence between integral points on a Mordell curve and the solutions to certain cubic Thue equations.

1.4. Reduction Theory

A reduction theory for a family of binary forms should consist of three things, a definition of what it means for a form to be reduced, a reduction algorithm which takes an arbitrary form and gives a properly equivalent reduced form, and a generating algorithm for producing all reduced forms up to equivalence with prescribed values for the generators of the invariant ring. That is, a reduction theory should describe a convenient family of forms, the reduced forms, to serve as a fundamental domain for proper equivelance of binary forms, and all computational methods required to work with this family of reduced forms.

The reduction theory of binary quadratic forms is classical, dating back to Gauss. A reduction theory for binary quartic forms was given by Birch and Swinnerton-Dyer in [17]. A reduction theory for binary cubic forms and an improved reduction theory for binary quartic forms was given by Cremona in [18]. However, we note that a small family of forms is not produced by the generating algorithm of these reduction theories, those whose reduced proper equivalent has vanishing leading coefficient. A convenient notion of reduced and a reduction algorithm was given for forms of higher degree by Julia in his treatise [19] although his definition is not explicit. More recent and explicit results are due to Cremona and Stoll in [20]. A generating algorithm is not known in degree five and higher.

In Chapter III we give a generating algorithm for the family of diagonalizable forms, Algorithm 3.5.1. Although generators of the invariant ring for all forms in arbitrary degree are not known, the invariant ring for diagonalizable forms is determined by the discriminant. Thus our algorithm instead produces all diagonalizable forms up to equivalence with a given discriminant.

An implementation of this algorithm in Sage can be found on the author's website:

`https://cdethier.github.io`.

We end Chapter III with some computational examples that were produced using this code. In particular, we have verified that Theorem 1.4 in [9] holds with $r = 5$, $h = 1$, and $m = 5$ if the assumption on the discriminant is dropped. The results of these computations can also be found on the author's website.

CHAPTER II

DIAGONALIZABLE QUARTIC THUE EQUATIONS WITH NEGATIVE DISCRIMINANT

2.1. Gap Principles

Suppose that $F(x,y) \in \mathbb{Z}[x,y]$ is a binary integral quartic form with resolvent forms ξ and η which satisfy (1.3). There are multiple choices for ξ and η, for example if ξ, η is one choice then $-\xi, i\eta$ is another. For the remainder of this chapter we fix a pair with real coefficients and define the corresponding scaled forms u and v so that $F = u^4 + v^4$ and (1.3) both hold. Again there are multiple ways to do this, so we fix a pair u and v with real coefficients.

We define

$$Z = Z(x,y) = \max\{|u(x,y)|, |v(x,y)|\}.$$

and

$$\zeta = \zeta(x,y) = \frac{|F(x,y)|}{Z^4(x,y)}.$$

When we are considering multiple solutions (x_i, y_i) indexed by i, for convenience we will frequently use the notation $\zeta_i = \zeta(x_i, y_i)$, $Z_i = Z(x_i, y_i)$, $\xi_i = \xi(x_i, y_i)$, etc. Furthermore, we will denote the solution to the inequality $0 < |F(x,y)| \leq h$ for which ζ is largest by (x_0, y_0). We also treat (x,y) and $(-x,-y)$ as the same solution, because Z only depends on $|u|$ and $|v|$.

The following is a result from [9], see the remark in that paper following Definition 5.3. We recall the proof here:

Lemma 2.1.1. *If* $|j| > 2\sqrt{h}$ *and the primitive integer pair* $(x_i, y_i) \neq (x_0, y_0)$ *satisfies* $0 < |F(x_i, y_i)| \leq h$*, then* $\zeta(x_i, y_i) < 1$.

Proof. Suppose to the contrary that $(x_i, y_i) \neq (x_0, y_0)$ is a solution to this equation with $\zeta_i \geq 1$. Then

$$u_0 v_i - u_i v_0 = (\alpha\delta - \beta\gamma)(x_0 y_i - x_i y_0) = j(x_0 y_i - x_i y_0) \neq 0.$$

From this we conclude that

$$|j| \leq |u_0 v_i| + |u_i v_0| \leq 2 Z_0 Z_i.$$

which we can use as follows:

$$|j| \leq 2 Z_0 Z_i = 2 \frac{|F_0|^{1/4}}{\zeta_0^{1/4}} \frac{|F_i|^{1/4}}{\zeta_i^{1/4}} \leq 2\sqrt{h}$$

because $\zeta_0, \zeta_i \geq 1$. It follows by contraposition that $|j| > 2\sqrt{h}$ and $(x_i, y_i) \neq (x_0, y_0)$, then $\zeta(x_i, y_i) < 1$. □

Suppose that ω is a fourth root of unity. For our fixed pair of resolvent forms η and ξ, we say that the solution (x, y) to $0 < |F(x, y)| \leq h$ is related to ω if

$$\left| \omega - \frac{\eta(x, y)}{\xi(x, y)} \right| = \min_{0 \leq k \leq 3} \left| e^{2k\pi i/4} - \frac{\eta(x, y)}{\xi(x, y)} \right|.$$

As ξ and η were assumed to have real coefficients, any solution must be related to one of the real fourth roots of unity.

Motivated by the previous lemma, we exclude the solution with largest ζ. We define S_ω to be the set of solutions related to ω, and S'_ω the collection of solutions

related to ω, excluding the solution whose ζ-value is largest. We index the elements of S'_ω as $(x_1, y_1), \ldots, (x_k, y_k)$ and once again adopt the notation Z_i, ζ_i, u_i, etc. Further, we may order the solutions in S'_ω to have decreasing ζ-values. That is, $\zeta_{i+1} \leq \zeta_i$ for all $1 \leq i \leq k-1$.

The following lemma originates in [11] and provides useful gap principles. We use the statements found in [9, Lemma 5.6] and [9, Lemma 5.7]

Lemma 2.1.2. *Assume that $|S'_\omega| \geq 2$ and $h < \frac{1}{4}|j|^2$. Let $(x_0, y_0) \in S'_\omega$ with largest ζ-value and $(x, y) \in S'_\omega$ a different solution. Then*

$$Z(x, y) \geq \frac{|j|}{2h^{1/4}}. \tag{2.1}$$

and

$$Z_i \geq \frac{|j|}{2h} Z_{i-1}^3. \tag{2.2}$$

Under the assumption $h < \frac{1}{4}|j|^2$, it follows that all elements of S'_ω have ζ-value less than 1 by Lemma 2.1.1, so we used that assumption rather than the assumption $\zeta_{i-1} < 1$ given in [9].

Lemma 2.1.3. *By convention, we label the elements of S'_ω as $(x_1, y_1), \ldots, (x_k, y_k)$ and order them by decreasing ζ-value. Suppose that $|S'_\omega| \geq 2$ and $h < \frac{1}{4}|j|^2$. Under these assumptions*

$$Z_k \geq \frac{|j|^{a_1(k)}}{2^{a_1(k)} h^{a_2(k)}}, \tag{2.3}$$

where the constants $a_1(k)$ and $a_2(k)$ are defined as follows:

$$a_1(k) := \frac{3^k - 1}{2} + 3^{k-1}$$
$$a_2(k) := \frac{3^k - 1}{2} + \frac{3^{k-1}}{4}.$$

Proof. We begin by applying (2.2) repeatedly to Z_k:

$$Z_k \geq \frac{|j|}{2h} Z_{k-1}^3 \geq \left(\frac{|j|}{2h}\right)^4 Z_{k-2}^9 \geq \ldots \geq \left(\frac{|j|}{2h}\right)^{b(k)} Z_1^{3^{k-1}},$$

where

$$b(k) = \sum_{i=0}^{k-1} 3^i = \frac{3^k - 1}{2}.$$

Finally, we apply (2.1) to Z_1 to obtain

$$Z_k \geq \left(\frac{|j|}{2h}\right)^{b(k)} \left(\frac{|j|}{2h^{1/4}}\right)^{3^{k-1}} = \frac{|j|^{b(k)+3^{k-1}}}{2^{b(k)+3^{k-1}} h^{b(k)+\frac{3^{k-1}}{4}}} = \frac{|j|^{a_1(k)}}{2^{a_1(k)} h^{a_2(k)}}.$$

□

2.2. Some Constants and Lemmas

Following [9], we define the constants $c_{n,g}$, $c_1(n,g)$ and $c_2(n,g)$ for $n \in \mathbb{N}$ and $g \in \{0,1\}$ as follows:

$$\begin{aligned} c_{n,g} &:= 4^n \left(12\sqrt{D}\right)^{n+g} \left(\frac{2}{\chi}\right)^{1-g} \\ c_1(n,g) &:= 2^{3n+2} |c_{n,g}| \\ c_2(n,g) &:= 2^{n+1-g} |c_{n,g}| \left(1 - \frac{2h}{Z_1^4}\right)^{-\frac{1}{2}(2n+1-g)} \frac{\left|\binom{n-g+1/4}{n+1-g}\binom{n-1/4}{n}\right|}{\binom{2n+1-g}{n}}. \end{aligned}$$

We state some bounds for $c_1(n,g)$ and $c_2(n,g)$ given in [9] which we will use:

$$|c_1(n,g)| \leq 2^{3n+2} 4^{2(2g+3n)+1} |j|^{2(g+n)+1} \tag{2.4}$$

$$|c_2(n,g)| \leq 2^{n+3} 4^{2(2g+3n)+1} |j|^{2(g+n)+1}. \tag{2.5}$$

These can be found in equations (60) and (61) in that paper.

We need some further results from [9] which explain the significance of $c_1(n,g)$ and $c_2(n,g)$. The following is [9, Lemma 7.3].

Lemma 2.2.1. *Let F be a diagonalizable binary quartic form. Let (x_1,y_1) and (x_2,y_2) be two solutions related to a fixed fourth root of unity, say ω, with $\zeta_2 \leq \zeta_1$. Assume that $Z_1^4 > 2h$ and $\Sigma_{n,g} \neq 0$. Then*

$$c_1(n,g)hZ_1^{4n+1-g}Z_2^{-3} + c_2(n,g)h^{2n+1-g}Z_1^{-4(n+1-g)+1-g}Z_2 > 1. \tag{2.6}$$

And this is Lemma 7.4 from that paper.

Lemma 2.2.2. *If $n \in \mathbb{N}$ and $I \in \{0,1\}$, then at most one of $\{\Sigma_{n,0}, \Sigma_{n+I,1}\}$ can vanish.*

2.3. Strengthening the Gap Principle

Throughout this section, we assume that S'_ω has k elements, indexed as $(x_1,y_1),\ldots,(x_k,y_k)$. Our aim is to show that under certain conditions this is a contradiction, in order to conclude that $|S'_\omega| \leq k-1$.

We begin by defining the constants C_i and E_i for $i = 0,1,2$. Throughout these definitions, $n \geq 2$ and $k \geq 3$. The E's are given as follows:

$$\begin{aligned} E_0(k) &:= \frac{4a_1(k-1)}{1+4a_2(k-1)} \\ E_1(k,g) &:= \frac{-2g+(4+g)a_1(k-1)}{4+(4+g)a_2(k-1)} \\ E_2(n,k,g) &:= \frac{-8n-14+2g+(8n-5+g)a_1(k-1)}{6n+4+(8n-5+g)a_2(k-1)} \end{aligned}$$

and the C's are given as $C_i = 2^{\Theta_i}$, where

$$\Theta_0 := \frac{-1 - 4a_1(k-1)}{1 + 4a_2(k-1)}$$
$$\Theta_1 := \frac{-24 - 8g - (4+g)a_1(k-1)}{4 + (4+g)a_2(k-1)}$$
$$\Theta_2 := \frac{3\log_2(3) - 54n - 66 - 8g - (8n-5+g)a_1(k-1)}{6n + 4 + (8n-5+g)a_2(k-1)}.$$

Lemma 2.3.1. *Suppose that $k \geq 3$ is fixed integer, and that h satisfies*

$$h < \frac{1}{4}|j|^2 \tag{2.7}$$

as well as

$$h \leq \min_{0 \leq i \leq 2} C_i |j|^{E_i} \tag{2.8}$$

for all $n \geq 2$ and $g = 0, 1$. Then

$$Z_k \geq (0.75)2^{-13n-13}|j|^{-2n-3}h^{-2n-1}Z_{k-1}^{4n} \tag{2.9}$$

for all $n \in \mathbb{N}$.

Proof. During this proof we will frequently use Lemma 2.2.1 applied to Z_{k-1} and Z_k. This Lemma requires the assumption that $Z_{k-1}^4 > 2h$. This is always the case, as

$$Z_{k-1}^4 \geq \left(\frac{|j|^{a_1(k-1)}}{2^{a_1(k-1)}h^{a_2(k-1)}}\right)^4 > 2h$$

using (2.3) and our assumption in (2.8) that $h < C_0(k)|j|^{E_0(k)}$.

This argument is a proof by induction. Beginning with the base case, $n = 1$, we cube (2.2) and rearrange to fit the first term of the left side of (2.6):

$$Z_k^3 \geq \left(\frac{|j|}{2h}\right)^3 Z_{k-1}^9$$

$$hc_1(1,g)Z_k^{-3}Z_{k-1}^{5-g} \leq c_1(1,g)|j|^{-3}2^3h^4Z_{k-1}^{-4-g}.$$

Now we apply (2.4) to $c_1(1,g)$ and (2.3) to Z_{k-1}:

$$\begin{aligned} hc_1(1,g)Z_k^{-3}Z_{k-1}^{5-g} &\leq h^4 2^3 |j|^{-3} \left(2^5 4^{4g+7}|j|^{2g+3}\right) \left(\frac{|j|^{a_1(k-1)}}{2^{a_1(k-1)}h^{a_2(k-1)}}\right)^{-4-g} \\ &= 2^{d_1}|j|^{d_2}h^{d_3}, \end{aligned}$$

where the exponents d_1, d_2, and d_3 are given as follows:

$$d_1 = 22 + 8g + (4+g)a_1(k-1)$$

$$d_2 = 2g - (4+g)a_1(k-1)$$

$$d_3 = 4 + (4+g)a_2(k-1).$$

Because of our assumption in (2.8) that $h < C_1(k,g)|j|^{E_1(k,g)}$, it follows that

$$c_1(1,g)hZ_k^{-3}Z_{k-1}^{5-g} < 0.25.$$

According to Lemma 2.2.2, $\Sigma_{1,0}$ and $\Sigma_{1,1}$ cannot both be zero. We choose whichever $\Sigma_{1,g}$ is nonzero and apply Lemma 2.2.1 to Z_k and Z_{k-1} to conclude that[1]

$$c_2(1,g)h^{3-g}Z_{k-1}^{3g-7}Z_k > 0.75.$$

Rearranging and applying (2.5) to $c_2(1,g)$, we see that

$$\begin{aligned} Z_k &> (0.75)2^{-18-8g}|j|^{-2g-3}h^{g-3}Z_{k-1}^{7-3g} \\ &\geq (0.75)2^{-26}|j|^{-5}h^{-3}Z_{k-1}^4 \end{aligned}$$

This last inequality required that $h \geq 1$, $|j| \geq 1$, which follows from $h < \frac{1}{4}|j|^2$ in (2.8), and $Z_{k-1} \geq 1$, which follows from (2.1) and $h < \frac{1}{4}|j|^2$. Since this is (2.9) with $n = 1$, this completes the base case.

We begin the induction argument by cubing the induction assumption and rearranging towards the first term of the left side of (2.6) with $n+1$:

$$\begin{aligned} Z_k^3 &\geq (0.75)^3 2^{-39n-39}|j|^{-6n-9}h^{-6n-3}Z_{k-1}^{12n} \\ hc_1(n+1,g)Z_k^{-3}Z_{k-1}^{4n+5} &\leq (0.75)^{-3}c_1(n+1,g)2^{39n+39}|j|^{6n+9}h^{6n+4}Z_{k-1}^{5-8n+g}. \end{aligned}$$

The left hand side is now the first term in (2.6), so we attempt to show that the right hand side is less than 0.25. To do this, we first make use of (2.4) applied to $c_1(n+1,g)$, then (2.3) applied to Z_{k-1}. Doing this second step requires the

[1]It is possible to make these arguments with 0.25 replaced by any $0 < \alpha < 1$. However, $\alpha = 0.25$ maximizes the expression $\alpha(1-\alpha)^3$ which appears in our C_2 constant.

assumption $h < \frac{1}{4}|j|^2$.

$$\begin{aligned} c_1(n+1,g)hZ_k^{-3}Z_{k-1}^{4n+5-g} &\leq (0.75)^{-3}2^{54n+8g+58}|j|^{8n+2g+12}h^{6n+4}Z_{k-1}^{-8n-4} \\ &\leq (0.75)^{-3}2^{d_4}|j|^{d_5}h^{d_6}, \end{aligned}$$

where the exponents d_4, d_5, and d_6 are given as follows:

$$\begin{aligned} d_4 &= 54n + 8g + 58 + (8n + g - 5)a_1(k-1) \\ d_5 &= 8n + 2g + 12 + (5 - 8n - g)a_1(k-1) \\ d_6 &= 6n + 4 + (8n + g - 5)a_2(k-1). \end{aligned}$$

By our assumption (2.8) that $h \leq C_2(n,k,g)|j|^{E_2(n,k,g)}$, it follows that

$$c_1(n+1,g)hZ_k^{-3}Z_{k-2}^{4n+5-g} < 0.25.$$

According to Lemma 2.2.2, $\Sigma_{n+1,0}$ and $\Sigma_{n+1,1}$ cannot both be zero. We choose whichever $\Sigma_{n+1,g}$ is nonzero and apply Lemma 2.2.1 to Z_k and Z_{k-1} to conclude that

$$c_2(n+1,g)h^{2n+3-g}Z_{k-1}^{-4n-7+3g}Z_k > 0.75.$$

Rearranging and applying (2.5), we see that

$$\begin{aligned} Z_k &> (0.75)2^{-13n-8g-18}|j|^{-2n-2g-3}h^{g-2n-3}Z_{k-1}^{4n+7-3g} \\ &\geq (0.75)2^{-13n-26}|j|^{-2n-5}h^{-2n-3}Z_{k-1}^{4n+4} \end{aligned}$$

Once again, we have used $h \geq 1$, $|j| \geq 1$ and $Z_{k-1} \geq 1$. These follow from $h < \frac{1}{4}|j|^2$ in (2.8) and (2.1). Since this is (2.9) with $n \to n+1$, we have completed the induction argument. □

We will show that the value of k for S'_ω leads to a contradiction. Before doing this, we first define two more constants, $E_3(k)$ given as follows:

$$E_3(k) := \frac{-2 + 4a_1(k-1)}{2 + 4a_2(k-1)}.$$

and $C_3(k)$ given as $C_3(k) = 2^{\Theta_3}$ where

$$\Theta_3 := \frac{-13 - 4a_1(k-1)}{2 + 4a_2(k-1)}.$$

Lemma 2.3.2. *Suppose that, in addition to* (2.7) *and* (2.8)*, we also assume that*

$$h < C_3(k)|j|^{E_3(k)}. \tag{2.10}$$

Then then inequality

$$0 < |F(x,y)| \leq h$$

has at most $2k$ *solutions.*

Proof. It suffices to show that under (2.10) that Lemma 2.3.1 leads to a contradiction, as we built that Lemma assuming that $|S'_\omega| = k$, and S'_ω contains all solutions related to a particular fourth root of unity except the one with largest ζ-value. As we have noted, solutions can only be related to two of the fourth roots of unity because u and v have real coefficients, as $I_F < 0$.

To derive a contradiction, we will show that the right side of (2.9) goes to ∞ as $n \to \infty$. To do this, we rearrange (2.10):

$$\begin{aligned} h &< C_3(k)|j|^{E_5(k)} \\ h^{2+4a_2(k-1)} &< 2^{13-4a_1(k-1)}|j|^{-2+4a_1(k-1)} \\ 1 &< 2^{-13}|j|^{-2}h^{-2}\left(\frac{|j|^{a_1(k-1)}}{2^{a_1(k-1)}h^{a_2(k-1)}}\right)^4 \\ 1 &< 2^{-13}|j|^{-2}h^{-2}Z_{k-1}^4. \end{aligned}$$

In the right side of (2.9) this quantity is being raised to the nth power, which will go to ∞. □

2.4. Reduction of Coefficients

Proof of Theorem 1.3.1. We complete the proof of Theorem 1.3.1 by comparing the constants C_i and E_i. We aim to show that for a fixed $k \geq 3$, $E_2(2,k,0)$ is minimal among the E_i and $C_2(2,k,0)$ is minimal among the C_i with $i = 0,1,2,3$, $n \geq 2$, and $g = 0,1$. This will show that $h < C_2(2,k,0)|j|^{E_2(2,k,0)}$ is the most restrictive constraint between (2.8) and (2.10), hence the only necessary one.

To do this we need to show several inqualities of the form $E_i(n_1,k,g_2) \leq E_j(n_2,k,g_2)$ and similar with the exponents of the C_i. This amounts to verifying several inequalities of the form

$$\frac{\xi_1 + \eta_1 a_1(k-1)}{\theta_1 \pm \eta_1 a_2(k-1)} \leq \frac{\xi_2 + \eta_2 a_1(k-1)}{\theta_2 \pm \eta_2 a_2(k-1)}. \tag{2.11}$$

Where $\pm$ is taken to be $+$ for the E's and $-$ for the C's. The constants ξ_1, η_1, θ_1, ξ_2, η_2, and θ_2 may depend on n or g, but not k. To do this, we clear denominators

and organize by the coefficients of 1, $a_1(k-1)$, and $a_2(k-1)$. Noticing that the a_1a_2 term always cancels, we define Φ as

$$\Phi = (\xi_2\theta_1 - \xi_1\theta_2) + (\eta_2\theta_1 - \eta_1\theta_2)a_1(k-1) + (\xi_2\eta_1 - \xi_1\eta_2)a_2(k-1)$$

and check that $\Phi \geq 0$ because this implies (2.11). For notation, we use $\Phi_{E,i}$ when checking the inequality $E_2(2,k,0) \leq E_i$ and $\Phi_{C,i}$ when checking the inequality $C_2(2,k,0) \leq C_i$. We use the notation $\ell = \log_2(3)$ and expand in terms of 3^k to obtain the following expressions for Φ:

$$\begin{aligned}
18\Phi_{E,0} &= 685 \cdot 3^k - 1017 \\
\tfrac{9}{2}\Phi_{E,1} &= 225 - 198g + (130 + 27g)3^k \\
\Phi_{E,2} &= 110 - 55n - 2g + (-842 + 421n + 131g)3^{k-2} \\
\tfrac{18}{7}\Phi_{E,3} &= -108 + 79 \cdot 3^k \\
36\Phi_{C,0} &= -5688 + 108\ell + (4265 - 841\ell)\,3^k \\
36\Phi_{C,1} &= 3816 - 216\ell - 5868g + 54\ell g + (2824 - 84\ell + 442g - 21\ell)\,3^k \\
36\Phi_{C,2} &= 13536 + 432\ell - 6768n - 216n\ell - 5868g + 54g\ell + \\
&\qquad + (-9932 + 336\ell + 4966n - 168n\ell + 442g - 21g\ell)\,3^k \\
\tfrac{36}{7}\Phi_{C,3} &= -598 + (493 - 12\ell)\,3^k.
\end{aligned}$$

One may verify that each of these expressions is non-negative, using the restrictions $n \geq 2$, $g = 0, 1$, and $k \geq 3$ where appropriate. □

2.5. Proof of Theorem 1.3.2, Forms with Small Discriminant

The method for this proof is to apply Theorem 1.3.1 with $h = 1$. The bounds on h in terms of j lead to upper bounds on Δ using $\Delta = -4^4 j^{12}$. These in turn lead to upper bounds on I_F using $27\Delta = 4I^3 - J^2$. We then find all forms F with $J_F = 0$ and I_F down to this bound and solve $|F(x, y)| = 1$ for each form.

Unfortunately, using $k = 3$ requires that we solve $|F(x, y)| = 1$ for all forms with (approximately)

$$0 > I_F > -2.4 \times 10^9.$$

which far exceeds our computational resources. Using $k = 4$ gives more reasonable bounds, (approximately)

$$0 > I_F > -2600.$$

Of course, this gives a weaker result. We see no reason why Theorem 1.3.2 should be false with eight replaced by six, but showing that statement is out of reach of our computational resources using this method.

Our presentation of these methods was inspired by [16].

Proof of Theorem 1.3.2. Applying Theorem 1.3.1 with $k = 4$ and $h = 1$ shows that $|F(x, y)| = 1$ has at most eight solutions for forms F with $I_F < -2593$. The remaining forms are handled by direct computation.

To find all such forms, we use an algorithm given by Cremona in [18, Section 4.6]. This algorithm misses the family of forms whose leading coefficient is zero when reduced. This issue is explicitly highlighed in [17] where Birch and Swinnerton-Dyer describe a similar algorithm.

These forms can be handled separately. If $F(x, y)$ has a leading coefficient of zero, then $F(x, y) = yC(x, y)$, where $C(x, y)$ is a cubic form. The equation

$yC(x,y) = \pm 1$ requires $y = \pm 1$ and $C(x,y) = \pm 1$ with the same sign, as y and $C(x,y)$ are both integers. Putting these together, we arrive at $C(x, \pm 1) = \pm 1$, which describes the roots of two cubic polynomials. Thus, $F(x,y) = \pm 1$ has at most six solutions.

Here we give a brief description of Cremona's algorithm for the case $I < 0$ and $J = 0$. To find all forms

$$F(x,y) = ax^4 + bx^3y + cx^2y^2 + dxy^3 + ey^4,$$

with $J_F = 0$ and a given negative value for I_F, we loop on a, b, and c using the bounds for a and b given by

$$|a| \leq \frac{2}{3\sqrt{3}}\sqrt{-I}$$
$$-2|a| < b \leq 2$$

and the bounds on c derived from the definition of the seminvariant H:

$$H = 8ac - 3b^2 \tag{2.12}$$

and the following bounds on H:

$$\max\left\{\frac{4}{3}I, -B_a\right\} \leq H \leq \min\{0, B_a\}$$

where B_a is given by

$$B_a = \frac{2}{3}\sqrt{-4I}\sqrt{-4I - 27a^2}.$$

These can be found in [18, Proposition 14]. Given a, b, and c one can find the seminvariant H using (2.12) and the seminvariant R using the identity

$$H^3 - 48Ia^2H + 64Ja^3 = -27R^2.$$

Then one can calculate d and e using the definition of R:

$$R = b^3 + 8a^2d - 4abc$$

and the definition of I:

$$I = 12ae - 3bd + c^2,$$

checking for integrality of R, d, and e after calculating each. Note that this algorithm is simplified by observing that when $J = 0$ it follows that I is divisible by three.

The results of these computations can be found on the author's website:

`https://cdethier.github.io`.

The file `forms.pdf` contains a list of forms with $J_F = 0$ and $0 > I_F > -3000$, organized in descending order of I_F. We claim that the list of forms in this pdf contains at least one form in each $SL_2(\mathbb{Z})$ orbit, however we do not claim that these forms are distinct up to $SL_2(\mathbb{Z})$ action.

Now that we have obtained a presentation of all forms of interest, we compute the solutions to $F(x, y) = 1$ and $F(x, y) = -1$ using PARI. The solutions of each equation are also given in the file `forms.pdf`. Table 1 lists the number of forms

TABLE 1 The number of Forms with a given number of solutions to the Thue equations $F(x,y) = 1$ and $F(x,y) = -1$.

$F(x,y) = 1$	$F(x,y) = -1$	# Forms
0	0	7346
0	1	1003
0	2	97
0	3	5
1	0	1003
1	1	146
1	2	3
2	0	97
2	1	3
3	0	5

with $J_F = 0$ and $0 > I_F > -3000$ with a given number of solutions to $F(x,y) = 1$ and $F(x,y) = -1$:

Crucially, none of these forms have more that eight primitive solutions to $|F(x,y)| = 1$, which completes the proof. □

We continue with some remarks about our computations. None of the forms discovered have more than three primitive solutions. These computations are consistent with observations, for example in [5], that most upper bounds for the number of solutions to a Thue equation are not sharp. Furthermore, all forms which have exactly three primitive solutions are diagonal, that is, they have shape $ax^4 + by^4$. The fact that these forms have more solutions than the rest is due to the fact that we count (x,y) and $(x,-y)$ as separate solutions, which we would not do when studying the family of diagonal forms in even degree.

This upper bound on the number of solutions with a diagonal form is not unexpected, see [21] for example. However, it is unexpected that this bound would hold for all quartic diagonalizable forms with negative discriminant.

2.6. Reduction of Elliptic Curves

Now we show how to bound the number of integral points on the elliptic curve

$$Y^2 = X^3 + NX = X(X^2 + N) \tag{1.8}$$

by bounding the number of solutions of a certain family of quartic Thue's inequalities. This reduction is due to Tzanakis and can be found in [15]. The case with $N < 0$ can be found in [9]. We recall it here to establish notation and to be self-contained.

Let N be a positive square-free integer. We consider the integral points on the elliptic curve (1.8). As X and $X^2 + NX$ are integers and Y^2 is a square integer, the square-free parts of X and $X^2 + N$ must be the same. Conversely, and X with X and X^2+N having identical square-free parts will lead to an integral point on (1.8). We will use the notation

$$X = dy^2, \text{ and } X^2 - N = dx^2.$$

From their definition x and y satisfy the equation $x^2 - dy^4 = \frac{N}{d}$. We may now focus on the quartic equation

$$X^2 - dY^4 = k, \tag{2.13}$$

where N and k are positive integers, and $d > 1$ is a positive square-free integer. Conversely, a solution to (2.13) also produces an integral point on (1.8) with $N = kd$.

Since it was assumed that N is square-free, the integer k is also square-free and is relatively prime to d. Let U_d be the number of solutions to equation (2.13).

Then the summation

$$\sum_{d|N} U_d \tag{2.14}$$

provides an upper bound for the number of solutions to (1.8). We calculate these upper bounds by counting integral solutions to the equation

$$X^2 - dY^2 = k \tag{2.15}$$

and detect those where Y is a square.

We begin by studying the structure of the solutions of this equation. Suppose that $(X, Y) \in \mathbb{Z}^2$ with $XY \neq 0$ is a solution to (2.15). Define

$$\alpha = X + Y\sqrt{d},$$

and for $i \in \mathbb{Z}$, define $X_i, Y_i \in \mathbb{Z}$ as follows:

$$X_i + Y_i\sqrt{d} = \alpha\epsilon_d^i$$

where ϵ_d is the fundamental unit in the order $\mathbb{Z}[\sqrt{d}]$.

Defined in this way, $(X_i, Y_i) \in \mathbb{Z}^2$ is also a solution to (2.15). We refer to the set of all such (X_i, Y_i) as the *class of solutions* of (2.15) associated to (X, Y). Walsh in [22] showed that there are at most 2^ω classes of solutions to (2.15) under the assumption that k is square-free and $D > 0$, see Corollary 3.1 in that paper. So we concern ourselves with bounding the number of solutions in a fixed class of solutions, C.

Let Y_0 be the least positive value of Y which occurs in C and let X_0 be the corresponding integer from C so that $X_0^2 - dY_0^2 = k$. We call $X_0 + Y_0\sqrt{d}$ the *fundamental solution* of the class C.

Now suppose that (X, Y) is a solution to (2.13), so that (X, Y^2) is a solution to (2.15). If $X_0 + Y_0\sqrt{d}$ is the fundmental solution of the class of solutions of $X + Y^2\sqrt{d}$, then

$$X + Y^2\sqrt{d} = \left(X_0 + Y_0\sqrt{d}\right)\epsilon_d^i \tag{2.16}$$

for some i. Then there are integers j, s, t such that

$$X + Y^2\sqrt{d} = \left(s + t\sqrt{d}\right)\epsilon_d^{2j} \tag{2.17}$$

by taking either

$$s + t\sqrt{d} = X_0 + Y_0\sqrt{d} \qquad \text{when } i \text{ is even, or}$$
$$s + t\sqrt{d} = \left(X_0 + Y_0\sqrt{d}\right)\epsilon_d \qquad \text{when } i \text{ is odd.}$$

Now suppose $\epsilon_d^j = m + n\sqrt{d}$. Then we have $m^2 - dn^2 = 1$ and expanding (2.17) we see that

$$Y^2 = tm^2 + 2smn + tDn^2.$$

Multiplying this identity by t, completing the square, and using the fact that $s^2 - dt^2 = k$, we obtain

$$-(tm + sn)^2 + kn^2 + tY^2 = 0. \tag{2.18}$$

The following is [15, Lemma].

Lemma 2.6.1. *Let a, b, c be nonzero integers with $\gcd(a, b, c) = 1$, and such that the equation*

$$a\mathfrak{X}^2 + b\mathfrak{Y}^2 + c\mathfrak{Z}^2 = 0 \tag{2.19}$$

has a solution in integers $\mathfrak{X}, \mathfrak{Y}, \mathfrak{Z}$ not all zero. Then there are integers R_1, S_1, T_1, R_2, S_2, T_2, and z_1 depending only on a, b, c satisfying the relations

$$R_1T_2 + R_2T_1 = 2S_1S_2,$$
$$S_2^2 - R_2T_2 = -acz_1^2$$
$$S_1^2 - R_1T_1 = -bcz_1^2$$

and a nonzero integer δ, also depending only on a, b, c such that for every nonzero solution $(\mathfrak{X}, \mathfrak{Y}, \mathfrak{Z})$ of (2.19), there exist integers Q, x, y, and a divisor P of δ so that

$$P\mathfrak{X} = Q(R_1x^2 - S_1xy + T_1y^2)$$
$$P\mathfrak{Y} = Q(R_2x^2 - 2S_2xy + T_2y^2).$$

Moreover if $\gcd(\mathfrak{X}, \mathfrak{Y}, \mathfrak{Z})$ is bounded, then an upper bound for Q can be found.

Furthermore, Walsh showed in [22] that the integers R_1, T_1, R_2, T_2 satisfy $R_1T_2 - R_2T_1 = 0$.

Applying Lemma 2.6.1 to (2.18) with $a = -1$, $b = k$, and $c = t$, we conclude that producing a solution to (2.18) is equivalent to producing a primitive solution to

$$F(u, v) = (Pt/Q)^2, \tag{2.20}$$

where $F(x,y) = A_1^2(x,y) - A_2^2(x,y)$ if we define A_1 and A_2 as

$$A_1(x,y) := (R_1 - sR_2)x^2 - 2(S_1 - sS_2)xy + (T_1 - sT_2)y^2$$
$$A_2(x,y) := R_2tx^2 - 2S_2txy + T_2ty^2.$$

We summarize some properties of this particular Thue equation in the following proposition:

Proposition 2.6.2. *Let $F(x,y)$ be the quartic form with coefficients given above. Then*

1. *$F(x,1)$ has exactly two real roots and no repeated roots,*
2. *$J_F = 0$,*
3. *$I_F = 48kt^3T_2R_2z_1^2d$,*
4. *$I_F < 0$.*

Proof. 1) Solving $F(x,1) = 0$ is equivalent to solving

$$A_1(x,1) = \pm\sqrt{d}A_2(x,1).$$

We make the substitution $w = s \pm t\sqrt{d}$, and this becomes

$$p(x) := (R_1 - wR_2)x^2 - 2(S_1 - wS_2)x + (T_1 - wT_2) = 0.$$

To check if the roots of this polynomial are real, we must check positivity of the discriminant of $p(x)$. We do this using the identities from Lemma 2.6.1.

$$\begin{aligned}\frac{1}{4}\Delta_p &= (S_1 - wS_2)^2 - (R_1 - wR_2)(T_1 - wT_2) \\ &= S_1^2 - 2wS_1S_2 + w^2S^2 - R_1T_1 + wR_1T_2 + wR_2T_1 - w^2R_2T_2 \\ &= -ktz_1^2 + wtz_1^2 \\ &= tz_1^2(w^2 - k).\end{aligned}$$

As t and z_1^2 are both positive, we must determine whether $w^2 - k$ is positive, negative, or zero:

$$\begin{aligned}w^2 - k &= S^2 \pm 2st\sqrt{d} + t^2d - s^2 + t^2d \\ &= 2t^2d \pm 2st\sqrt{d} \\ &= 2t\sqrt{d}(t\sqrt{d} \pm s).\end{aligned}$$

Now we must determine whether $t\sqrt{d} \pm s$ is positive, negative, or zero. To do this, we note that

$$(s + t\sqrt{d})(-s + t\sqrt{d}) = -s^2 + dt^2 = -k < 0,$$

which implies that exactly one of $s + t\sqrt{d}$ and $-s + t\sqrt{d}$ is negative, the other is positive, and neither are zero. In fact, $-s + t\sqrt{d} < 0$ as $s, t > 0$. Thus we see that $F(x, 1)$ has two real roots and two non-real roots, as well as no repeated roots.

2) is proved in [22], while 3) is shown in [9].

4) It follows from 1) that $\Delta_F < 0$, which implies that $I_F < 0$ from the identity $27\Delta_F = 4I_F^3 - J_F^2$

□

2.7. Bombieri-Schmidt Reduction

Proposition 2.7.1. *Let $\mathfrak{G}$ be the set of quartic forms $F(x,y) \in \mathbb{Z}[x,y]$ that are irreducible over $\mathbb{Q}$ with $I_F < 0$ and $J_F = 0$. Let $\mathfrak{N}$ be an upper bound for the number of solutions of quartic Thue equations*

$$F(x,y) = 1$$

as F varies over the elements of $\mathfrak{G}$. Then for $h \in \mathbb{N}$ and $G(x,y) \in \mathfrak{G}$, the equation

$$G(x,y) = h \tag{2.21}$$

has at most

$$\mathfrak{N}4^{\omega(h)}$$

primitive solutions.

Proof. This is a special case of [14, Lemma 7]. In that proof (2.21) is reduced to certain other Thue equations with other forms in $\mathfrak{G}$ by reducing $G(x,y)$ through the action of certain matrices from $\mathrm{GL}_2(\mathbb{Z})$. These new forms will have $J_F = 0$ because applying this action to a diagonalized form clearly yields a diagonalized form. Furthermore, the matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

will act on a root α by

$$\alpha \mapsto \frac{a \cdot \alpha + b}{c \cdot \alpha + d}.$$

From this it is clear that real roots will map to real roots and nonreal roots will map to nonreal roots. Thus these new forms will also have $I < 0$. □

2.8. Proof of Theorem 1.3.4

Proof of Theorem 1.3.4. Tracing back through our reduction of the elliptic curve, the number of integral points on (1.8) is at most

$$\sum_{d|N} U_d,$$

where U_d is an upper bound for the number of solutions to the Thue equation (2.13). Every two of these solutions is derived from one solution to (2.15) as Y is squared. The solutions to (2.15) split into classes of solutions. As k is square-free, Walsh showed in [22] that there are at most $2^{\omega(k)}$ such classes. The number of solutions in each class is the number of solutions to the quartic Thue equation (2.20), which is at most $8 \cdot 4^{\omega\left(P^2t^2/Q^2\right)}$, applying Theorem 1.3.3. Akhtari in [9] gives the following upper bound for $\omega(P^2t^2/Q^2)$ (see the proof of Corollary 5.1):

$$\omega\left(\frac{P^2t^2}{Q^2}\right) \leq 2 + \frac{\log\left(\epsilon_d^{3/2}\sqrt{|K|/2d}\right)}{\log 4}$$

where $K = N/d$. Hence it follows that (1.8) has at most

$$\sum_{d|N} U_d \leq 2^{15/2}\sqrt{N}\sum_{d|N}\frac{2^{\omega(N/d)}\epsilon_d^{3/2}}{d}$$

integral points. □

CHAPTER III

COMPUTATIONAL THEORY OF DIAGONALIZABLE FORMS

3.1. Proof of Theorem 1.2.1

Proof. If F is a diagonalizable form of degree r, then it follows from the computation in (1.4) that $H_F = (L_1(x,y)L_2(x,y))^{r-2}$, where L_1 and L_2 are linear forms proportional to the resolvent forms of F, $a\alpha x + \beta y$ and $\gamma x + \delta y$. If L_1 and L_2 were proportional, then the resolvent forms of F would be proportional, implying that $\Delta_F = 0$, which contradicts the assumption $\Delta_F \neq 0$ for diagonalizable forms.

Conversely, suppose that $F(x,y) \in \mathbb{Z}[x,y]$ is a binary integral form with nonzero discriminant whose Hessian H_F is the $r-2$ power of a quadratic form with non-proportional linear factors. We will prove that F is diagonalizable. Suppose that the two linear factors are $\xi(x,y)$ and $\eta(x,y)$, given by

$$\xi(x,y) = \alpha x + \beta y \qquad \text{and} \qquad \eta(x,y) = \gamma x + \delta y.$$

Since ξ and η are not proportional, one may write

$$x = p\xi + q\eta$$
$$y = s\xi + t\eta$$

for some $p, q, s, t \in \mathbb{C}$. In fact, these can be given explicitly by inverting the implicit linear substitution matrix. We use $a_0, a_1, \ldots, a_r$ for the coefficients of F in ξ and η,

and Φ the form F viewed with this perspective:

$$\begin{aligned} F(x,y) &= F(p\xi + q\eta, s\xi + t\eta) \\ &= a_0\xi^r + a_1\xi^{r-1}\eta + \ldots + a_{r-1}\xi\eta^{r-1} + a_r\eta^r \\ &= \Phi(\xi,\eta). \end{aligned}$$

We use $A_0, A_1, \ldots, A_{2r-2}$ for the coefficients of the Hessian of F in ξ and η. We note that it satisfies

$$\begin{aligned} H_\Phi(\xi,\eta) &= A_0\xi^{2r-2} + A_1\xi^{2r-3}\eta + \ldots + A_{2r-3}\xi\eta^{2r-3} + A_{2r-2}\eta^{2r-2} \\ &= (pt-sq)^2 H_F(x,y) = (pt-sq)^2(\xi\eta)^{r-2}, \end{aligned}$$

as the Hessian is a degree two covariant. Thus we conclude that

$$A_0 = \ldots = A_{r-3} = A_{r-1} = \ldots = A_r = 0 \qquad \text{and} \qquad A_{r-2} = (pt-sq)^2.$$

We will use explicit calculation of the Hessian coefficients of Φ (the A_k) in terms of the coefficients of Φ (the a_m) to show that $a_1 = a_2 = \ldots = a_{r-1} = 0$. This will show that F is diagonalizable. We begin with

$$\Phi(\xi,\eta) = \sum_{i=0}^{r} a_i\xi^{r-i}\eta^i,$$

from which we calculate the second order partial derivatives of Φ:

$$\begin{aligned}
\frac{\partial^2\Phi}{\partial\xi^2} &= \sum_{i=0}^{r-2}(r-i)(r-i-1)a_i\xi^{r-i-2}\eta^i \\
\frac{\partial^2\Phi}{\partial\eta^2} &= \sum_{i=2}^{r}(i)(i-1)a_i\xi^{r-i}\eta^{i-2} = \sum_{i=0}^{r-2}(i+2)(i+1)a_{i+2}\xi^{r-i-2}\eta^i \\
\frac{\partial^2\Phi}{\partial\xi\eta} &= \sum_{i=1}^{r-1}(r-i)(i)a_i\xi^{r-i-1}\eta^{i-1} = \sum_{i=0}^{r-2}(r-i-1)(i+1)a_{i+1}\xi^{r-i-2}\eta^i.
\end{aligned}$$

We have reindexed these sums so that each summand has the same power of ξ and η. This leaves us with the following expression for the H_Φ:

$$\begin{aligned}
H_\Phi = &\left(\sum_{i=0}^{r-2}(r-i)(r-i-1)a_i\xi^{r-i-2}\eta^i\right)\left(\sum_{j=0}^{r-2}(j+2)(j+1)a_{j+2}\xi^{r-j-2}\eta^j\right) \\
&-\left(\sum_{i=0}^{r-2}(r-i-1)(i+1)a_{i+1}\xi^{r-i-2}\eta^i\right)^2.
\end{aligned}$$

We now collect terms by the resulting powers of ξ and η, because A_k is the coefficient of η^k in H_Φ. This yields the following identity:

$$\begin{aligned}
A_k = \sum_{\substack{i+j=k \\ 0\le i,j\le r-2}} &[(r-i)(r-i-1)(j+1)(j+2)a_ia_{j+2} \\
&-(r-i-1)(i+1)(r-j-1)(j+1)a_{i+1}a_{j+1}].
\end{aligned} \tag{3.1}$$

It will be convenient for the rest of this argument to assume that $a_0 \neq 0$ or $a_r \neq 0$. First we assume that $a_0 = 0$ and $a_r = 0$ to arrive at a contradiction. First we assume r is odd, then (3.1) with $k = 0$ is

$$0 = 2r(r-1)a_0a_2 - (r-1)^2a_1^2$$

which shows that $a_0 = 0$ forces $a_1 = 0$. Using $a_0 = a_1 = 0$ in (3.1) with $k = 2$ is

$$0 = 12r(r-1)a_0a_4 + (6r-6)a_1a_3 + 4(r-2)^2a_2^2 + 2(r-2)(r-3)a_2^2$$

which shows that $a_2 = 0$. One may inductively show that $a_m = 0$ using the identity $A_{2m-2} = 0$ up to $2m - 2 = r - 3$ (one must be sure that the coefficient of a_m^2 is nonzero — this can be checked explicitly in general). Similarly, (3.1) with $k = 2r-4$ shows that $a_r = 0$ forces $a_{r-1} = 0$. One may also inductively build downwards to show that $a_m = 0$ using $A_{2m-2} = 0$ down to $2m - 2 = r - 1$ (again, checking that the coefficient of a_m is nonzero). Thus, if r is odd and $a_0 = a_r = 0$, then $a_m = 0$ for $0 \leq m \leq r$. This means $F = 0$, so $H_F = 0$ which contradicts our assumptions.

If r is even, the same argument applies. However, it fails to show that $a_{r/2} = 0$, as $A_{r-2} \neq 0$. This implies that F has the form $F(x, y) = a_{r/2}(\xi\eta)^{r/2}$. However, this form has $\Delta_F = 0$, so it is ignored by our assumptions.

So we may thus assume that $a_0 \neq 0$ or $a_r \neq 0$. As (3.1) is symmetric under the permutation $m \mapsto r - m$ of the subscripts of the a_m, we will only make our argument in the case where $a_0 \neq 0$. The case where $a_r \neq 0$ can be argued symmetrically.

We will use (3.1) to solve for successive values of a_m in terms of a_1 and a_0, starting with solving $A_0 = 0$ for a_2, and proceeding inductively by solving $A_k = 0$ for a_{k+2}. We claim that this leads to the following presentation of a_m for $2 \leq m \leq r - 1$:

$$a_m = \frac{(r-1)\dots(r-m+1)}{r^{m-1}m!}\frac{a_1^m}{a_0^{m-1}}. \tag{3.2}$$

Before arguing this claim, we begin by showing that if $a_i, a_{i+1}, a_{j+1}, a_{j+2}$ all have this presentation, then the i, j term of the sum in (3.1) is equal to 0. This will be

useful in the proof of (3.2) and after. To show the i, j, term of the sum is 0, we compute both expressions directly. Here is the first:

$$\begin{aligned}
&(r-i)(r-i-1)(j+1)(j+1)a_i a_{j+2} = \\
&= (r-i)(r-i-1)(j+1)(j+2)\frac{(r-1)\dots(r-i+1)}{r^{i-1}i!}\frac{a_1^i}{a_0^{i-1}}\cdot \\
&\quad\cdot\frac{(r-1)\dots(r-j-1)}{r^{j+1}(j+2)!}\frac{a_1^{j+2}}{a_0^{j+1}} \\
&= \frac{(r-1)\dots(r-i-1)(r-1)\dots(r-j-1)}{r^{i+j}i!j!}\frac{a_1^{i+j+2}}{a_0^{i+j}}.
\end{aligned}$$

And the second:

$$\begin{aligned}
&-(r-i-1)(i+1)(r-j-1)(j+1)a_{i+1}a_{j+1} = \\
&= -(r-i-1)(i+1)(r-j-1)(j+1)\frac{(r-1)\dots(r-i)}{r^i(i+1)!}\frac{a_1^{i+1}}{a_0^i}\cdot \\
&\quad\cdot\frac{(r-1)\dots(r-j)}{n^j(j+1)!}\frac{a_1^{j+1}}{a_0^j} \\
&= \frac{-(r-1)\dots(r-j-1)(r-1)\dots(r-j-1)}{r^{i+j}i!j!}\frac{a_1^{i+j+2}}{a_0^{i+j}}.
\end{aligned}$$

These expressions cancel, showing that the i, j term of the sum is 0, provided $a_2, \dots, a_{k+1}$ have the presentation given in (3.2).

Now we can show that (3.2) is the case. Of course we will argue this claim by induction. For the base case, $A_0 = 0$ is the identity

$$2r(r-1)a_0a_2 - (r-1)^2a_1^2 = 0,$$

which we solve to obtain

$$a_2 = \frac{r-1}{2r}\frac{a_1^2}{a_0},$$

which is (3.2) when $m = 2$.

For the induction step, suppose that we have used the equations up to $A_{k-1} = 0$ to solve for $a_1, \ldots, a_{k+1}$ and we are now proceeding to solve $A_k = 0$ for a_{k+2}. The first term of the sum in (3.1), the term with $i = 0$, contains the only value of a_m which we have not solved for, a_{k+2}. The remaining terms of the sum contain only values of a_m that we have solved for. Hence, as we have shown, all terms except the first are 0. Now that the first term of the sum is the only one remaining, we may solve for a_{k+2}:

$$\begin{aligned} r(k+2)a_0a_{k+2} &= (r-k-1)a_1a_{k-1} \\ r(k+2)a_0a_{k+2} &= (r-k-1)a_1\frac{(r-1)\ldots(r-k)}{r^k(k+1)!}\frac{a_1^{k+1}}{a_1^k} \\ a_{k+2} &= \frac{(r-1)\ldots(r-k-1)}{r^{k+1}(k+2)!}\frac{a_1^{k+2}}{a_0^{k+1}}. \end{aligned}$$

Hence by induction the a_m all have this presentation. We may continue this induction argument up to solving for a_{r-1} in the equation $A_{r-3} = 0$.

Now we assume that $a_1 \neq 0$ and hope to reach a contradiction. We skip $k = r-2$ for the moment and proceed to solve $A_{r-1} = 0$ for a_r. As we have shown, all terms with values of a_m given in (3.2) are 0. The only remaining term is the one containing a_r, which we use to solve for a_r. Note that this requires $a_1 \neq 0$ for cancellation:

$$\begin{aligned} r(r-1)(r-2)a_1a_r &= 2a_2a_{r-1} \\ r(r-1)(r-2)a_1a_r &= 2\frac{r-1}{2r}\frac{a_1^2}{a_0}\frac{(r-1)\ldots(2)}{r^{r-2}(r-1)!}\frac{a_1^{r-1}}{a_0^{r-2}} \\ a_r &= \frac{1}{r^r}\frac{a_1^r}{a_0^{r-1}}. \end{aligned}$$

This we have shown that a_m has the shape given in (3.2). However, moving back to $A_{r-2} = (pt - sq)^2$ shows that $(pt - sq)^2 = 0$ as every a_m in (3.1) has the shape given in (3.2).

Hence we have contradicted the assumption that $Q(x, y)$ has non-proportional linear factors, showing that in fact we must have $a_1 = 0$. Examining (3.2), we see that in fact $a_1 = \ldots = a_{r-1} = 0$. This shows that F is diagonalizable. □

3.2. Preliminaries

This section further introduces the notation of diagonalizable forms and covers some known results concerning the coefficients which are required for our reduction theory. Where possible we have used the notation of [9].

A diagonalizable form may also be presented as

$$F(x, y) = \alpha_1(x - \beta_1 y)^r - \gamma_1(x - \delta_1 y)^r, \tag{3.3}$$

with the corresponding restriction

$$j^r = \alpha_1\gamma_1(\delta_1 - \beta_1)^r \neq 0.$$

If F is to have integral (or rational) coefficients, then $\alpha, \beta, \gamma, \delta$ or alternatively $\alpha_1, \beta_1, \gamma_1, \delta_1$ must satisfy certain algebraic conditions. The following lemma gives such conditions for (3.3).

Lemma 3.2.1. *Suppose that F is a diagonalizable form with rational coefficients that has been diagonalized as in* (3.3)*. Then one of the following must be the case:*

a) $\alpha_1, \beta_1, \gamma_1, \delta_1 \in \mathbb{Q}$

b) $[\mathbb{Q}(\beta_1) : \mathbb{Q}] = 2$ *and* δ_1 *is the algebraic conjugate of* β_1*. Furthermore* $\alpha_1, -\gamma_1 \in \mathbb{Q}(\beta_1)$ *and are also algebraic conjugates.*

Suppose that $F(x, y)$ *has integral coefficients and* $\mathbb{Q}(\beta_1) = \mathbb{Q}(\sqrt{D})$*. Let* $\mathcal{O} = \mathcal{O}_{\mathbb{Q}(\sqrt{D})}$ *be the ring of integers in* $\mathbb{Q}(\sqrt{D})$*. Then the coefficients of*

$$r(r-1)\sqrt{D}\alpha_1(x-\beta_1 y)^r \qquad \textit{and} \qquad r(r-1)\sqrt{D}\gamma_1(x-\delta_1 y)^r$$

are in $\mathcal{O}$*. In particular,* $r(r-1)\sqrt{D}\alpha_1, r(r-1)\sqrt{D}\gamma_1 \in \mathcal{O}$*.*

The first part of this lemma is due to Voutier, and can be found as Lemma 4.1 in [23]. The second part of this lemma is due to Akhtari, Saradha, and Sharma, and can be found as Lemma 3.2 in [9].

We also note that j is similar to the discriminant of F, which we notate $\Delta = \Delta_F$. Explicitly,

$$\Delta = (-1)^{\frac{(r-1)(r-2)}{2}} r^r j^{r(r-1)}. \tag{3.4}$$

This can be found as equation (17) in [9]. Furthermore, there is a constant $\chi \in \mathbb{C}$ such that

$$(\alpha x + \beta y)(\gamma x + \delta y) = \chi(Ax^2 + Bxy + Cy^2), \tag{3.5}$$

where $A, B, C \in \mathbb{Z}$. The discriminant of this integral quadratic form we will call D,

$$D = D_F = B^2 - 4AC.$$

We may further assume that $\gcd(A, B, C) = 1$ as otherwise their greatest common divisor could be included in χ. This assumption ensures that A, B, C, D, and χ are uniquely defined. To be explicitly clear, we follow this convention even in the special case that two of A, B, and C are zero, which ensures that the third is ± 1.

Computing the quadratic discriminant of both sides of (3.5) gives

$$j^2 = \chi^2 D. \tag{3.6}$$

Choosing arbitrary rth roots of α_1 and γ_1, we see that

$$\chi(Ax^2 + Bxy + Cy^2) = \alpha_1^{1/r}\gamma_1^{1/r}(x - \beta_1)(x - \delta_1), \tag{3.7}$$

it follows that $\beta_1, \delta_1 \in \mathbb{Q}(\sqrt{D})$ which justifies our repeated use of D from Lemma 3.2.1. We note that all of this information can be found in [9].

The Hessian of F is a covariant of F defined as

$$H(x, y) = \left(\frac{\partial^2 F}{\partial x^2}\right)\left(\frac{\partial^2 F}{\partial y^2}\right) - \left(\frac{\partial^2 F}{\partial xy}\right)^2.$$

When defined this way, it is clear that the Hessian of a binary integral form will itself have integral coefficients. Computing the Hessian of the diagonalizable form

$$F(x, y) = (\alpha x + \beta y)^r - (\gamma x + \delta y)^r \tag{1.2}$$

and using the definition of χ in (3.5), we see that

$$H = -r^2(r - 1)^2 j^2 \chi^{r-2}(Ax^2 + Bxy + Cy^2)^{r-2}. \tag{3.8}$$

As A, B, and C are relatively prime and $(Ax^2 + Bxy + Cy^2)^{r-2}$ includes terms with coefficients A^{r-2}, B^{r-2}, and C^{r-2}, we must have

$$r^2(r - 1)^2 j^2 \chi^{r-2} \in \mathbb{Z} \tag{3.9}$$

for the Hessian of F to have integral coefficients. This observation is originally due to Gauss, as noted by Siegel in [11].

3.3. The Discriminant of a Diagonalizable Form

Suppose that $F(x, y) \in \mathbb{Z}[x, y]$ is a diagonalizable form of degree r. Not every integer is a possible value for the discriminant Δ_F. One can give restrictions on the value of Δ_F which we discuss in this section.

Suppose that r is even, we consider the expression

$$r^2(r-1)^2\chi^{r-2}j^2D^{(r-2)/2} = r^2(r-1)^2j^r. \tag{3.10}$$

As $D \in \mathbb{Z}$, it follows from (3.9) that

$$r^2(r-1)^2j^r \in \mathbb{Z}. \tag{3.11}$$

When r is odd, we consider the expression

$$r^4(r-1)^4\chi^{2r-2}j^4D^{r-2} = r^4(r-1)^4j^{2r}. \tag{3.12}$$

Again, as $D \in \mathbb{Z}$, it follows that from (3.9) that

$$r^4(r-1)^4j^{2r} \in \mathbb{Z}. \tag{3.13}$$

From (3.11) and (3.13) we build the following result, which is a useful necessary condition for diagonalizability.

Lemma 3.3.1. *Suppose that F is a diagonalizable form of degree $r \geq 4$ with integral coefficients. Then $\chi^r \in \mathbb{Q}$.*

When r is even there is an integer $\mathcal{D}$ such that $\Delta = r\mathcal{D}^{r-1}$. In addition, j must satisfy $rj^r \in \mathbb{Z}$. In fact, $\mathcal{D}$ can be taken to be

$$\mathcal{D} = (-1)^{(r+2)/2} r j^r,$$

and conversely these two identities uniquely determine $\mathcal{D}$ and j^r for a given Δ.

When r is odd there is an integer $\mathcal{D}$ such that $\Delta^2 = r^2\mathcal{D}^{r-1}$. In addition, j must satisfy $r^2 j^{2r} \in \mathbb{Z}$. In fact, $\mathcal{D}$ can be taken to satisfy

$$|\mathcal{D}| = \left| r^2 j^{2r} \right|,$$

and conversely these two identities uniquely determine $\mathcal{D}$ and j^{2r} up to sign for a given Δ. Furthermore j^{2r} and D have the same sign. If $r \equiv 1 \mod 4$ then $\Delta > 0$ and if $r \equiv 3 \mod 4$ then Δ and D have opposite signs.

Proof. We will verify each of these statements in order. For the first statement, it follows from (3.7) that

$$-(\chi A)^r = -\alpha_1 \gamma_1.$$

According to Lemma 3.2.1, there are two possibilities. If $\alpha_1, \gamma_1 \in \mathbb{Q}$, then $\chi^r \in \mathbb{Q}$ as $A \in \mathbb{Z}$. If $[\mathbb{Q}(\beta_1) : \mathbb{Q}] = 2$ with $\alpha_1, -\gamma_1$ conjugates in $\mathbb{Q}(\beta_1)$, then

$$-(\chi A)^r = -\alpha_1 \gamma_1 = \mathrm{Nm}(\alpha_1) \in \mathbb{Q},$$

from which it follows that $\chi^r \in \mathbb{Q}$.

We proceed with r even. We define $\mathcal{D}$ by

$$\mathcal{D} = (-1)^{(r+2)/2} r j^r. \tag{3.14}$$

It follows from (3.11) that $\mathcal{D} \in \mathbb{Q}$. It also follows from (3.4) that $\Delta = rm\mathcal{D}^{r-1}$. Let $\mathcal{D} = p/q$ with $p, q \in \mathbb{Z}$ and $q > 0$. Then $\Delta = r\mathcal{D}^{r-1} \in \mathbb{Z}$ implies that $q^{r-1}|r$. However, $q \geq 2$ and $2^{r-1} > r$ for $r \geq 4$ give a contradiction, showing that $\mathcal{D} \in \mathbb{Z}$.

The proof when r is odd is similar. We again define $\mathcal{D}$ by

$$\mathcal{D} = r^2 j^{2r}. \tag{3.15}$$

It follows from (3.13) that $\mathcal{D} \in \mathbb{Q}$. It also follows from (3.4) that $\Delta^2 = r^2\mathcal{D}^{r-1}$. Let $\mathcal{D} = p/q$ with $p, q \in \mathbb{Z}$, and $q > 0$. Then $\Delta^2 = r^2\mathcal{D}^{r-1} \in \mathbb{Z}$ implies that $q^{r-1}|r^2$. As r is odd, we must have $q \geq 3$. However, $3^{r-1} > r^2$ for $r \geq 5$, so it must be the case that $\mathcal{D} \in \mathbb{Z}$.

For the final statements, we note that when r is odd (3.13) shows that $j^{2r} \in \mathbb{Q}$. Furthermore, taking rth powers of (3.6) gives $j^{2r} = \chi^{2r} D^r$ which shows that j^{2r} and D have the same sign, as χ^r is rational and r is odd.

If $r \equiv 1 \mod 4$, then (3.4) becomes

$$\Delta = r^r (j^{2r})^{(r-1)/2},$$

from which we conclude that $\Delta > 0$ as $(r-1)/2$ is even. When $r \equiv 3 \mod 4$, then (3.4) becomes

$$\Delta = -r^r (j^{2r})^{(r-1)/2},$$

from which we conclude that Δ and D have opposite signs, as $(r-1)/2$ is odd and j^{2r} and D have the same sign. $\square$

Lemma 3.3.2. *Suppose that F is a diagonalizable form of degree r. Then $D = 1$ if and only if F is properly equivalent to a diagonal form, see* (1.7).

Suppose r is even. If $r(r-1)^2\mathcal{D}$ is not divisible by any $(r-2)/2$ powers, then F is diagonal. Furthermore, if F is diagonal with coefficients as in (1.7)*, then*

$$\mathcal{D} = (-1)^{r/2} rab.$$

Suppose r is odd. If $r^2(r-1)^4\mathcal{D}$ is not divisibile by any $r-2$ powers, then F is diagonal. Furthermore, if F is diagonal with coefficients as in (1.7)*, then*

$$\mathcal{D} = r^2 a^2 b^2.$$

Proof. For the very first statement, if F is a diagonal form, then (3.5) shows that $A = C = 0$. By our convention in the definitions of χ and D, we have $B = \pm 1$, hence $D = 1$. Conversely, suppose that $D = 1$. By the classical reduction theory of quadratic forms, if D is square, then the collection of quadratic forms given by

$$Q(x, y) = Ax^2 + Bxy,$$

where $B = \pm\sqrt{D}$ and $0 \leq A < |B|$ is a set of representatives for the family of quadratic forms with discriminant D up to equivalence. As $D = 1$ we may assume up to equivalence that $A = 0$ and $B = \pm 1$. However, examining (3.5) shows that this is clearly only possible if $\alpha = \delta = 0$ or if $\beta = \gamma = 0$. In either case F is diagonal.

Suppose that r is even. Then according to (3.10) and (3.11), it follows that $D^{(r-2)/2}$ divides $r^2(r-1)^2 j^r$. According to Lemma 3.3.1 however, $\mathcal{D} = (-1)^{(r+2)/2} r j^r$. Combining these shows that $D^{(r-2)/2}$ divides $r(r-1)^2\mathcal{D}$, so if $r(r-1)^2\mathcal{D}$ is free of $(r-2)/2$ powers, then $D = 1$ implying that F is equivalence to a diagonal form by the first statement. The proof when r is odd is similar, using the corresponding identities for when r is odd.

The remaining two statements about the discriminant of a diagonal form may be computed directly from the identities relating Δ and $\mathcal{D}$ in Lemma 3.3.1 and the following well-known identity for the discriminant of $f \in \mathbb{Z}[x]$,

$$\Delta_f = (-1)^{r(r-1)/2} a_0^{-1} \mathrm{Res}\,(f, f')$$

where Res indicates the resultant of two polynomials, and a_0 is the leading coefficient of f. □

3.4. Reduction Lemmas

Lemma 3.4.1. *Suppose that D is a positive non-square integer, and $u = a_1 + b_1\sqrt{D}$ is a unit in $\mathcal{O}_{\mathbb{Q}(\sqrt{D})}$. Let*

$$u^n = a_n + b_n\sqrt{D}.$$

Then for any $m \in \mathbb{Z}$, there is an $s \in \mathbb{Z}_{\geq 1}$ such that $a_s, b_s \in \mathbb{Z}$ and $m | b_s$.

Proof. We consider the sequences a_n and b_n. These sequences can be defined recursively, in the sense that a_n and b_n satisfy the following pair of recursive

identities:

$$a_n = a_{n-1}a_0 + b_{n-1}b_0D$$
$$b_n = a_{n-1}b_0 + b_{n-1}a_0,$$

as well as the following pair of descending recursive identities:

$$a_{n-1} = a_n a_0 - b_n b_0 D$$
$$b_{n-1} = -a_n b_0 + b_n a_0.$$

When we say these pairs of recursive identities are inverse, we mean that replacing a_{n-1} and b_{n-1} in the first pair of identities with the second pair of identities yields $a_n = a_n$ and $b_n = b_n$. Specifically,

$$a_n = (a_n a_0 - b_n b_0 D)a_0 + (-a_n b_0 + b_n a_0)b_0 D = a_n(a_0^2 - b_0^2 D) = a_n$$
$$b_n = (a_n a_0 - b_n b_0 D)b_0 + (-a_n b_0 + b_n a_0)a_0 = b_n(a_0^2 - b_0^2 D) = b_n,$$

as $a_0 + b_0\sqrt{D}$ is a unit.

These sequences take values in $\frac{1}{2}\mathbb{Z}$, so we may consider their values in the additive group $\frac{1}{2}\mathbb{Z}/\mathbb{Z}$. (Note that, although these recursive sequences are defined with multiplication which $\frac{1}{2}\mathbb{Z}$ isn't closed under, we may still reduce the values of the sequence modulo m.) Suppose that a_n and b_n have values $[a_n]$ and $[b_r]$ in the group $\frac{1}{2}\mathbb{Z}/\mathbb{Z}$. We define $c_n = ([a_n], [b_n])$. Using our relations, c_n can be calculated from c_{n-1} as well as from c_{n+1}.

As $c_n = (a_n \mod m, b_n \mod m)$ can only take on $(2m)^2$ possible values, c_n eventually contains a repeat, after which point it is periodic. Then, as c_{n-1} is

determined by c_n, we deduce that c_n is not just eventually periodic, but completely periodic. Because $c_0 = (1,0)$, there must be another power $s \in \mathbb{Z}_{\geq 0}$ such that $c_s = (1,0)$. For this s we have $a_s \in \mathbb{Z}$ and $b_s \in \mathbb{Z}$ with $m|b_s$. □

Lemma 3.4.2. *Suppose that we have a diagonalizable form F which has been diagonalized as*

$$F(x,y) = \alpha_1(x-\beta_1)^r - \gamma_1(x-\delta_1)^r.$$

Further suppose that for this F, one has that D is a positive non-square integer and $u = a_0 + b_0\sqrt{D}$ is the fundamental unit in $\mathbb{Q}(\sqrt{D})$. We adopt the notation

$$F_n(x,y) = u^n\alpha_1(x-\beta_1 y)^r - u^{-n}\gamma_1(x-\delta_1 y)^r.$$

Then there is a power $s \in \mathbb{Z}_{\geq 0}$ of u such that the forms F_n and F_{n+s} are properly equivalent for all $n \geq 0$.

Proof. To show this, we produce an $\mathrm{SL}_2(\mathbb{Z})$ matrix and an $s' \in \mathbb{Z}_{\geq 0}$ which has the effect of multiplying the linear forms of F by $u^{s'}$. That is, we show that there are $a,b,c,d \in \mathbb{Z}$ and an $s' \in \mathbb{Z}$ such that the substitution

$$x = aX + bY$$
$$y = cX + dY$$

satisfies $ad - bc = 1$ and yields

$$x - \beta_1 y = u^{s'}(X - \beta_1 Y)$$
$$x - \delta_1 y = u^{-s'}(X - \delta_1 Y).$$

Then the statement will follow for $s = rs'$. We may further assume that u has positive norm, as replacing u by u^2 may be accomplished by doubling s'.

Before producing this substitution, we must establish some notation. We know from Lemma 3.2.1 that $\beta_1, \delta_1 \in \mathbb{Q}(\sqrt{D})$. We also know from (3.7) that β_1 and δ_1 are algebraic conjugates, and the roots of $Ax^2 + Bxy + Cy^2$. So there are $\mu, \nu \in \mathbb{Q}$ such that

$$\beta_1 = \mu + \nu\sqrt{D}$$
$$\delta_1 = \mu - \nu\sqrt{D},$$

and $2A\mu, 2A\nu \in \mathbb{Z}$ by the quadratic formula. So there are $\mu', \nu' \in \mathbb{Z}$ such that $\mu = \frac{\mu'}{2A}$ and $\nu = \frac{\nu'}{2A}$.

We choose s' using Lemma 3.4.1. It follows from this Lemma that there is an $s' \in \mathbb{Z}_{\geq 0}$ such that $u^{s'} = a_{s'} + b_{s'}\sqrt{D}$ satisfies $a_{s'}, b_{s'} \in \mathbb{Z}$ as well as $2A\nu' | b_{s'}$. For our linear substitution we use the matrix

$$\begin{pmatrix} a_{s'} - \dfrac{b_{s'}\mu}{\nu} & \dfrac{b_{s'}\mu^2}{\nu} - b_{s'}\nu D \\ \dfrac{-b_{s'}}{\nu} & \dfrac{b_{s'}\mu}{\nu} + a_{s'} \end{pmatrix}. \tag{3.16}$$

This matrix has determinant one as u is a unit with positive norm:

$$\left(a_{s'} - \frac{b_{s'}\mu}{\nu}\right)\left(\frac{b_{s'}\mu}{\nu} + a_{s'}\right) - \left(\frac{b_{s'}\mu^2}{\nu} - b_{s'}\nu D\right)\left(\frac{-b_{s'}}{\nu}\right)$$
$$= (a_{s'})^2 - \left(\frac{b_{s'}\mu}{\nu}\right)^2 + \left(\frac{b_{s'}\mu}{\nu}\right)^2 - (b_{s'})^2 D = 1.$$

Furthermore, this matrix has integer entries. To see this,

$$a_{s'}, \frac{b_{s'}}{\nu} = \frac{2Ab_{s'}}{\nu'}, \frac{b_{s'}\mu}{\nu} = \frac{b_{s'}\mu'}{\nu'}, \frac{b_{s'}\mu^2}{\nu^2} = \frac{b_{s'}(\mu')^2}{2A\nu'} \in \mathbb{Z}$$

which all follow from the conditions on $a_{s'}$ and $b_{s'}$ guarranteed to us by Lemma 3.4.1. This shows that our substitution represents an equivalence of binary forms.

Finally, this matrix represents multiplication by $u^{s'}$ and $u^{-s'}$:

$$\begin{aligned}
x - \beta_1 y &= \left(a_{s'} - \frac{b_{s'}\mu}{\nu}\right) X + \left(\frac{b_{s'}\mu^2}{\nu} - b_{s'}\nu D\right) Y \\
&\quad - \left(\mu + \nu\sqrt{D}\right) \left[\left(\frac{-b_{s'}}{\nu}\right) X + \left(\frac{b_{s'}\mu}{\nu} + a_{s'}\right) Y\right] \\
&= \left(a_{s'} + b_{s'}\sqrt{D}\right) X - \left(a_{s'} + b_{s'}\sqrt{D}\right) \left(\mu + \nu\sqrt{D}\right) Y \\
&= u^{s'} \left(X - \beta_1 Y\right) \\
x - \gamma_1 y &= \left(a_{s'} - \frac{b_{s'}\mu}{\nu}\right) X + \left(\frac{b_{s'}\mu^2}{\nu} - b_{s'}\nu D\right) Y \\
&\quad - \left(\mu - \nu\sqrt{D}\right) \left[\left(\frac{-b_{s'}}{\nu}\right) X + \left(\frac{b_{s'}\mu}{\nu} + a_{s'}\right) Y\right] \\
&= \left(a_{s'} - b_{s'}\sqrt{D}\right) X - \left(a_{s'} - b_{s'}\sqrt{D}\right) \left(\mu - \nu\sqrt{D}\right) Y \\
&= u^{-s'} \left(X - \gamma_1 Y\right).
\end{aligned}$$

This completes the proof. □

3.5. Reduction Theory

This section gives an algorithm for producing all diagonalizable forms up to (proper) equivalence with a given discriminant. A summary of this algorithm is given in Algorithm 3.5.1. One should first find $\mathcal{D}$ using the identities with Δ given

in Lemma 3.3.1. Note that when r is odd $\mathcal{D}$ is assumed to be positive. One should now find all diagonal forms (if any) with this $\mathcal{D}$. Working towards a diagonal form

$$F(x, y) = ax^r + by^r,$$

one must simply use the identities relating $\mathcal{D}$, a, and b given in Lemma 3.3.2 to find a complete list of possibilities for a and b. Note that when r is even the following forms are properly equivalent:

$$ax^r + by^r \sim bx^r + ay^r.$$

When r is odd the following forms are properly equivalent:

$$ax^r + by^r \sim -bx^r + ay^r \sim -ax^r - by^r \sim bx^r - ay^r.$$

Now that we have obtained the family of diagonal forms, we proceed to find the forms which are not properly equivalent to a diagonal form. One should follow the identities in Lemma 3.3.1 to find j^r when r is even, and j^{2r} when r is odd. Note that when r is odd, j^{2r} is only determined up to sign by these identities, so one must proceed with both possibilities.

Next one should find a list of possibilities for D. In doing this, treat r even and r odd separately. When r is even, note from (3.10), (3.11), and Lemma 3.3.1 that $D^{(r-2)/2}$ divides $r(r-1)^2\mathcal{D}$. When r is odd, we note from (3.12), (3.13), and Lemma 3.3.1 that D^{r-2} is a divisor of $r^2(r-1)^4\mathcal{D}$. This gives a complete list of possibilities for D. Furthermore, one may ignore the case $D = 1$ as such forms are properly equivalent to diagonal forms by Lemma 3.3.2, and one may assume that

$D \equiv 0, 1 \mod 4$ as $D = B^2 - 4AC$. Finally, when r is odd Lemma 3.3.1 gives information on the sign of D in terms of the sign of j^{2n}. However, for r even one must consider positive and negative values for D.

One should next solve for χ^r by taking $j^2 = \chi^2 D$ to the $r/2$ or r depending on the parity of r. The specific values of χ and j are not needed for this algorithm. In fact, j^r or j^{2r} is not even needed beyond this step. This simplifies calculations greatly, as χ^r is rational and j^r or j^{2r} is rational while χ and j are not necessarily real.

Given D nonzero, one can produce a finite list of quadratic forms $Q(x, y) = Ax^2 + Bxy + Cy^2$ up to equivalence with discriminant D. As the case when D is square is not often included in the literature, we simply note that the list of forms $F(x, y) = Ax^2 + Bxy$ with $B = \pm\sqrt{D}$ and $0 \leq A < |B|$ suffices. Furthermore, for square D one may ignore the possibility $A = 0$, as such forms are properly equivalent to diagonal forms.

By (3.7), we may take β_1 and δ_1 to be the roots of Q. In the search for a list of possibilities for α_1 and γ_1, there are now three meaningful cases which one must treat separately. Ordered by complexity, they are:

i) D is square with $A \neq 0$.

ii) $D < 0$.

iii) $D > 0$ and D is not a perfect square.

We note that D is square with $A = 0$ does not need to be considered, as this would give one $A = C = 0$, hence a diagonal form.

First we consider the case when D is a square but $A \neq 0$. It follows from Lemma 3.2.1 that α_1 and γ_1 are rational, and $r(r-1)\sqrt{D}\alpha_1$ and $r(r-1)\sqrt{D}\gamma_1$

are integral. Furthermore, it follows from (3.7) that $\alpha_1\gamma_1 = (\chi A)^r$, from which we conclude that the divisors of

$$r^2(r-1)^2 D\alpha_1\gamma_1 = r^2(r-1)^2 D(\chi A)^r \in \mathbb{Z}$$

gives a complete list of possibilities for $n(n-1)\sqrt{D}\alpha_1$ and $r(r-1)\sqrt{D}\gamma_1$.

In the two remaining cases, it follows from Lemma 3.2.1 that $\alpha_1, -\gamma_1$ are conjugates in $\mathbb{Q}(\sqrt{D})$, and that

$$r(r-1)\sqrt{D}\alpha, r(r-1)\sqrt{D}\gamma_1 \in \mathcal{O}_{\mathbb{Q}(\sqrt{D})}.$$

It follows that $\sqrt{D}\alpha_1$ and $\sqrt{D}\gamma_1$ are conjugates, and hence that

$$\mathrm{Nm}(r(r-1)\sqrt{D}\alpha_1) = r^2(r-1)^2 D\alpha_1\gamma_1 = r^2(r-1)^2 D(\chi A)^r \in \mathbb{Z}.$$

A complete list of possibilities for $n(n-1)\sqrt{D}\alpha_1$ can thus be found by searching for a complete list of integers in $\mathcal{O}_{\mathbb{Q}(\sqrt{D})}$ with the norm given above. Fortunately, PARI has the function

```
bnfisintnorm(bnfinit(x^2 - D),N)
```

which gives the elements of $\mathcal{O}_{\mathbb{Q}(\sqrt{D})}$ which have norm N up to multiplication by units in $\mathcal{O}_{\mathbb{Q}(\sqrt{D})}$. If one would like to do this manually, the Lagrange–Matthews–Mollin algorithm may be applied. This algorithm is described in [24].

When $D < 0$, the unit group of $\mathcal{O}_{\mathbb{Q}(\sqrt{D})}$ is $\{\pm 1\}$, which, along with this PARI command, allows one to give a complete list of possibilities for $r(r-1)\sqrt{D}\alpha_1$. Taking the conjugate of α_1 then gives $-\gamma_1$.

In the remaining case, when $D < 0$, the unit group of $\mathcal{O}_{\mathbb{Q}(\sqrt{D})}$ is generated by -1 and some fundamental unit u, which can be found using either PARI or Sage. To find a complete list of possibilities for $r(r-1)\sqrt{D}\alpha_1$, one must therefore only indicate the highest power s of the fundamental unit necessary. Then multiplying the output of the PARI code by $\pm u^t$ for $0 \leq t \leq s$ gives a full list of possibilities for $r(r-1)\sqrt{D}\alpha_1$. To do this, find the smallest positive integer s' which makes the matrix given in (3.16) have integer coordinates and determinant $+1$. Then $rs' - 1$ is the largest power of u that one must consider.

Finally, one should check that all forms generated have integer coefficients and the appropriate discriminant. Multiplication by units when $D < 0$ in the previous step frequently yields forms with incorrect discriminant.

This algorithm is summarized in the following:

Algorithm 3.5.1. To find an exhaustive list of diagonalizable forms with a given discriminant, follow the following algorithm. For further explanations and justifications for any of these steps, we refer the reader to the preceding exposition.

(1) Find $\mathcal{D}$ using Lemma 3.3.1.

(2) Find all diagonal forms using Lemma 3.3.2.

(3) If r is even find j^r, and if r is odd find j^{2r}, both using Lemma 3.3.1.

(4) When r is even $D^{(r-2)/2}$ is a divisor of the integer expression $r(r-1)^2\mathcal{D}$. When r is odd D^{r-2} is a divisor of the integral expression $r^2(r-1)^4\mathcal{D}$. Ignore $D = 1$ in both cases.

(5) Find χ^r using $j^2 = \chi^2 D$.

(6) Find all possible quadratic forms $Q(x, y) = Ax^2 + Bxy + Cy^2$ with discriminant D using classical quadratic reduction theory. For square D ignore $A = 0$.

(7) The roots of Q are β_1 and δ_1.

(8) If D is square then $r(r-1)\sqrt{D}\alpha_1$ and $r(r-1)\sqrt{D}\gamma_1$ are integers whose product is $r^2(r-1)^2 D\chi^r A^r$, completing the algorithm for these D.

(9) Otherwise, use the PARI code

```
bnfisintnorm(bnfinit(x^2 - D),N)
```

with $N = r^2(r-1)^2 D(\chi A)^r$.

(10) If $D < 0$, multiply the results of this PARI code by ± 1 to obtain a complete list of possibilities for $r(r-1)\sqrt{D}\alpha_1$. Then $-\gamma_1$ is the conjugate of α_1, completing the algorithm for these D.

(11) If $D > 0$, find the smallest positive integer s' for which the matrix (3.16) has integer coordinates and determinant $+1$. Multiplying the output of the PARI code by $\pm u^t$ for $0 \leq t < rs'$, where u is the fundamental unit in $\mathbb{Q}(\sqrt{D})$, gives a complete list of possibilities for $r(r-1)\sqrt{D}\alpha_1$. Then $-\gamma_1$ is the conjugate of α_1, completing the algorithm for these D.

(12) Check that all forms have integer coefficients and the correct discriminant.

We finish with some remarks which can reduce the computation time necessary for running this algorithm. First, if the discriminant does not satisfy the conditions in Lemma 3.3.1, then there are are no diagonalizable forms with that discriminant.

Furthermore, in step 2, diagonal forms are not possible if $r \nmid \mathcal{D}$ when r is even, or $r^2 \nmid \mathcal{D}$ when r is odd by Lemma 3.3.2.

When finding D, we note that as

$$D = B^2 - 4AC$$

for integers A, B, and C, reducing modulo four implies that $D \equiv 0, 1 \mod 4$. Furthermore, according to Lemma 3.3.1 if $r \equiv 3 \mod 4$, then the sign of D is opposite that of the discriminant and if $r \equiv 1 \mod 4$, then the sign of D is the same as that of j^{2r}.

One may require the the coefficients of $Q = Ax^2 + Bxy + Cy^2$ are relatively prime, as a common factor in 3.5 could be included in χ. In particular, this means that if D is square and $A = 0$, one only need consider the quadratics xy and $-xy$.

When one is carrying out this algorithm for a large number of discriminants in the same degree, it can help to handle values of D which come up often in Algorithm 3.5.1. This is exemplified by Lemma 3.3.2 and Lemma 3.6.1.

Finally, as most applications involve Thue equations, we note that the number of solutions to the equation $F(x, y) = h$ with $x, y \in \mathbb{Z}$ does not change if F is replaced by an equivalent form. Hence for such applications, one does not need to verify that the matrix in (3.16) has determinant $+1$. In addition, one ignore $B < 0$ when producing a list of quadratic forms with discriminant D. Lastly, the following diagonal forms are equivalent when r is odd for every choice in each $\pm$,

$$\pm ax^r \pm by^r \sim \pm bx^r \pm ay^r.$$

3.6. Computational Example

In this section, we use our algorithm to verify that a special case of one theorem in [9] holds when the assumption on the size of the discriminant is removed. Before proceeding with this computation, we prove a Lemma.

Lemma 3.6.1. *Suppose that F is a quintic diagonalizable form. If $D_F = 4$ then F is equivalent to the form*

$$\left(\frac{a}{2} + \frac{b}{2}\right) x^5 + 5bx^4y + 20bx^3y^2 + 40bx^2y^3 + 40bxy^4 + 16by^5$$

for some choice of $a, b \in \mathbb{Z}$. This form has

$$\mathcal{D} = 25600a^2b^2.$$

If $D_F = -4$ then F is properly equivalent to the form

$$ax^5 - 5bx^4y - 10ax^3y^2 + 10bx^2y^3 + 5axy^4 - by^5$$

for some choice of $a, b \in \mathbb{Z}$. This form has

$$\mathcal{D} = 1600(a^2 + b^2)^2.$$

If $D_F = 8$ then F is properly equivalent to one of the following forms:

$$\begin{aligned} F_1(x,y) &= ax^5 + (5a+5b)x^4y + (30a+20b)x^3y^2 + (70a+50b)x^2y^3 \\ &\quad + (85a+60b)xy^4 + (41a+29b)y^5 \\ F_2(x,y) &= ax^5 - (5a-5b)x^4y + (30a-20b)x^3y^2 - (70a-50b)x^2y^3 \\ &\quad + (85a-60b)xy^4 - (41a+29b)y^5 \end{aligned}$$

for some choice of $a, b \in \mathbb{Z}$. Both of these forms have

$$\mathcal{D} = 12800(2a^2 - b^2)^2.$$

If $D_F = -8$ then F is properly equivalent to the form

$$ax^5 - 5bx^4 - 20ax^3 + 20bx^2 + 20ax - 4b$$

for some choice of $a, b \in \mathbb{Z}$. This form has

$$\mathcal{D} = 12800(2a^2 + b^2)^2.$$

Proof. To obtain each of these, we used the classical reduction theory of quadratic forms to produce a list of possibilities for $Ax^2 + Bxy + Cy^2$. From this we obtained β_1 and δ_1. We then let $\alpha_1' = a' + b'\sqrt{d}$ and $-\gamma_1' = a' - b'\sqrt{d}$, where d is the square-free part of D. We then expanded

$$\alpha_1'(x - \beta_1)^5 - \gamma_1'(x - \delta_1)^5$$

and used integrality of the coefficients to obtain bounds of the denominators of a' and b'. Using these denominators, we let a and b be the numerators of a' and b', then let $\alpha_1 = a + b\sqrt{d}$ and $-\gamma_1 = a - b\sqrt{d}$, and finally expanded

$$\alpha_1(x - \beta_1)^5 - \gamma_1(x - \delta_1)^5.$$

Improper equivalence is stated as otherwise there other forms with $D = 4$. □

The purpose of this Lemma is to demonstrate that certain values of D which commonly as a possibility in Algorithm 3.5.1 with $n = 5$ in fact do not lead to diagonalizable forms very frequently. Thus one may frequently ignore these possibilities.

We used Algorithm 3.5.1 to produce a complete list of quintic diagonalizable forms up to improper equivalence with $\mathcal{D} < 255137$. We were able to ignore several choices of D for most values of m using Lemmas 3.3.2 and 3.6.1. We proceeded to solve the Thue equation $|F(x, y)| = 1$ for each form, and considered the solutions (x, y) and $(-x, -y)$ as the same. The results of these computations can be found in the file `QuinticForms.pdf` on the author's website:

`https://cdethier.github.io/research.html`.

Crucially, none of these equations have more than four solutions. Thus we have verified that Theorem 1.4 in [9] holds with $r = 5$, $m = 5$, and $h = 1$ in the indefinite case if the assumption on the size of the discriminant is removed. Checking this theorem for $r = 5$ and $h = 1$ with $m = 3$ or $m = 4$ appears to be out of computational reach. For example, $m = 4$ would require one to check approximately all forms with $\mathcal{D} < 1.05 \times 10^9$.

REFERENCES CITED

[1] J. J. Sylvester. An essay on canonical forms: supplement to a sketch of a memoir on elimination, transformation and canonical forms. *George Bell, London*, 1851. Reprinted as Paper 34, pp. 203–216 in *The collected mathematical papers of James Joseph Sylvester, I: 1837–1853*, Cambridge University Press, 2012.

[2] J. J. Sylvester. On a remarkable discovery in the theory of canonical forms and of hyperdeterminants. *Philosophical Magazine*, 2:391–410, 1851. Reprinted as Paper 41, pp. 265–283 in *The collected mathematical papers of James Joseph Sylvester, I: 1837–1853*, Cambridge University Press, 2012.

[3] S. Gundelfinger. Zur Theorie der binären Formen. *J. Reine Angew. Math*, pages 413–424, 1886.

[4] P. J. Olver. *Classical invariant theory*. Cambridge University Press, 2003.

[5] S. Akhtari. The method of Thue–Siegel for binary quartic forms. *Acta Arithmetica*, 141(1):1–31, 2010.

[6] A. Thue. Über Annäherungswerte algebraischer Zahlen. *Journal für die reine und angewandte Mathematik*, 135:284–305, 1909.

[7] A. Thue. Berechnung aller Lösungen gewisser Gleichungen von der form $ax^r - by^r = f$. *Vid. Skrifter I Mat.-Naturv. Klasse*, pages 1–9, 1918.

[8] C. L. Siegel. Die Gleichung $ax^n - by^n = c$. *Mathematische Annalen*, 114:57–68, 1937.

[9] S. Akhtari, N. Saradha, and D. Sharma. Thue's inequalities and the hypergeometric method. *Ramanujan J.*, 45(2):521–567, 2018.

[10] S. Akhtari. Integral points on a certain family of elliptic curves. *J. Théor. Nombres Bordeaux*, 27(2):353–373, 2015.

[11] C. L. Siegel. Einige Erläuterungen zu Thues Untersuchungen über Annäherungs-werte algebraischer Zahlen und diophantische Gleichungen. *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II*, pages 169–195, 1970.

[12] I. Wakabayashi. On a family of quartic Thue inequalities. I. *J. Number Theory*, 66(1):70–84, 1997.

[13] I. Wakabayashi. On a family of quartic Thue inequalities. II. *J. Number Theory*, 80(1):60–88, 2000.

[14] E. Bombieri and W. M. Schmidt. On Thue's equation. *Inventiones mathematicae*, 88:69–82, 1987.

[15] N. Tzanakis. On the Diophantine equation $x^2 - dy^4 = k$. *Acta Arithmetica*, 46 (3):257–269, 1986.

[16] M. A. Bennett and A. Ghadermarzi. Mordell's equation: a classical approach. *LMS Journal of Computation and Mathematics*, 18:633–646, 2015.

[17] B. J. Birch and H. P. F. Swinnerton-Dyer. Notes on elliptic curves. I. *Journal für die reine und angewandte Mathematik*, 212:7–25, 1963.

[18] J. E. Cremona. Reduction of binary cubic and quartic forms. *LMS J. Comput. Math.*, 2:64–94, 1999.

[19] G. Julia. Études sur les formes binaires non quadratiques à indéterminées réelles ou complexes. *Mémoires de l'Académie des Sciences de l'Institut de France*, 55, 1917. Also published in *Œuvres de Gaston Julia* vol. 5 (1968–70).

[20] J. E. Cremona and M. Stoll. On the reduction theory of binary forms. *J. Reine Angew. Math.*, 565:79–99, 2003.

[21] M. A. Bennett and B. M. M. De Weger. On the Diophantine equation $|ax^n - by^n| = 1$. *Mathematics of Computation*, 67(221):413–438, 1998.

[22] P. G. Walsh. On the number of large integer points on elliptic curves. *Acta Arithmetica*, 138(4):317–327, 2009.

[23] P. M. Voutier. Thue's fundamentaltheorem, I: The general case. *Acta Arithmetica*, 143:101–144, 2010.

[24] K. Matthews. The Diophantine equation $x^2 - Dy^2 = N$, $D > 0$. *Expositiones Mathematicae*, 18:323–332, 2000.